

Food supplements and their efficacy

Citation for published version (APA):

Kleijnen, J. M. P. (1991). *Food supplements and their efficacy*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19911128jk>

Document status and date:

Published: 01/01/1991

DOI:

[10.26481/dis.19911128jk](https://doi.org/10.26481/dis.19911128jk)

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.

FOOD SUPPLEMENTS AND THEIR EFFICACY

PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus, Prof. Mr. M.J. Cohen
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen
op donderdag 28 november 1991 om 14.00 uur

door

Joseph Maria Pieter Kleijnen

Promotor: Prof. dr. P.G. Knipschild

Becoordelingscommissie: Prof. dr. J.A. Knottnerus (voorzitter)
Prof. dr. F. ten Hoor
Prof. dr. G.J. Kok
Prof. dr. H. Philipsen
Dr. A.L.M. Verbeek, KU Nijmegen

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Kleijnen, Joseph Maria Pieter

Food supplements and their efficacy / Joseph Maria Pieter

Kleijnen ; [ill. by the author]. - [S.l. : s.n.]. - Ill.

Thesis Maastricht. - With ref. - With summary in Dutch.

ISBN 90-9004581-3

NUGI 747

Subject headings: food supplements ; efficacy / meta-analysis.

Lay-out: Thum Aarts Maastricht

Productie: Datawyse Maastricht

Omslagontwerp: Jan Eggen

Druk: Datawyse Maastricht / Krips Repro Meppel

De publicatie van dit proefschrift werd mede mogelijk gemaakt door financiële steun van:
Stichting Dr. Ir. J.H.J. van de Laar, VSM Geneesmiddelen B.V., Biohorma Beheer B.V.,
Lichtwer Pharma GmbH, Pharmafood B.V., Indros & Pflüger B.V., Dr. Willmar Schwabe
GmbH & Co.

CONTENTS

	<u>page</u>
Contents	3
Chapter 1: General introduction	5
Chapter 2: Searching the literature with on-line computer databases	9
Chapter 3: Belief in the efficacy of alternative medicine	19
Chapter 4: Vitamin C and the common cold	21
Chapter 5: Vitamin E and cardiovascular disease	29
Chapter 6: Niacin and vitamin B6 in mental functioning	33
Chapter 7: Vitamin B6 and the premenstrual syndrome	45
Chapter 8: Evening primrose oil	51
Chapter 9: Garlic and cardiovascular disease	63
Chapter 10: Garlic supplements and cardiovascular disease	73
Chapter 11: Ginkgo biloba	83
Chapter 12: Clinical trials of homoeopathy	95
Chapter 13: Publication bias	105
Chapter 14: Epilogue	115
Summary	119
Samenvatting	123
Curriculum vitae	127
Dankwoord	127

CHAPTER 1

GENERAL INTRODUCTION

Research into the efficacy of food supplements

This dissertation contains reviews on the efficacy of food supplements. People take food supplements because most of them expect that an additional intake of these products may have some beneficial health effect. Food supplements are substances which are normally present in the diet, often in limited amounts. A strict definition of food supplements would include all vitamins, many minerals and trace elements and numerous other products containing at least some substances which can be found in a normal diet. A broader definition would also include "natural remedies" such as evening primrose oil, ginkgo biloba and homoeopathy. Some of these supplements have become quite popular. In the Netherlands, for example, many people take supplemental vitamin C, multi-vitamins or garlic.¹ Homoeopathy is an outlook in medicine developed by Samuel Hahnemann from the similia concept: "similia similibus curentur" which implies that a diluted, "potentised" agent, which (when undiluted) in healthy individuals induces complaints resembling those of the patient, can be used to cure the patient.

This work is part of a project sponsored by the Dutch Ministry of Welfare, Public Health and Cultural Affairs: "Review of the literature on the efficacy of food supplements and alternative medicine."

In this project several other subjects have been addressed. In an early stage a review on the effects of ginseng in elderly people with vitality problems was published,² as well as an empirical study of iridology as a diagnostic aid.^{3,4} Also, a series of articles on the efficacy of acupuncture,^{5,6} an article about chelation therapy (EDTA) for cardiovascular disease,⁹ and an article on placebo-effects in efficacy research have been published.¹⁰ Presently, amid 1991, new subjects such as procaine therapy against the effects of ageing and alternative treatments in patients with cancer are being evaluated.

There are several related projects at our department of Epidemiology & Biostatistics. For instance, the evidence of the effects of manipulative therapies for musculoskeletal problems is being assessed.^{11,12} In another project, the relation between the consumption of food supplements, for example garlic, and the incidence of cancer is being investigated. In a randomized clinical trial, the effects of a megadose of vitamin C in the treatment of pressure sores will be assessed. Also in a randomized clinical trial, the effects of Dabao, a Chinese hair restorer containing many herbs, have been evaluated.¹³ Finally, the relation between beta-carotene and cancer has been reviewed, and in a trial it was shown that beta-carotene has no beneficial effects on the regression of cervical dysplasia.^{14,15}

Reviews of the literature

When the question arises about the efficacy of a certain intervention, such as a food supplement, there are two approaches for obtaining an answer. Many people immediately think of performing controlled trials, but often the results of such trials can already be found in the medical literature. Nowadays, in the era of "publish or perish", it is very hard to keep up with all new developments, especially when one is all day busy treating patients. It may very well be that many questions have already been investigated. Before considering to start a time and money consuming trial, it may be worthwhile to review the evidence that is already there. Assessment of the evidence may lead to ideas for new

trials on the same subject, which would add something new to the already existing evidence. This is exactly what this dissertation is about.

The classical way in which a review of the literature was performed may be, simplified, described as follows. An eminent specialist in the field of interest has collected several studies, the results of some of these are presented, and a conclusion is drawn. In the recent past, there has been an increasing demand for a more objective approach: a description of how the search of the literature is performed, a detailed motivation of the inclusion criteria for trials; description and motivation of the way in which the evidence is assessed and how the conclusions are drawn. Also, the technique of statistical pooling of the results of many trials, in an attempt to increase the power of the evidence, has become popular.

This more scientific approach, with or without statistical pooling, is called meta-analysis. Reviews in which a methodological assessment of the quality of trials is emphasized, are called criteria-based meta-analyses. This approach stresses that it makes no sense to combine evidence of poor quality with evidence from well performed trials. A substantiated selection is made of high quality trials, and the conclusions are only based on their evidence.

Another problem which obtains increasing attention in meta-analyses is publication bias. Publication bias occurs if the results from studies which have not been published differ from the published ones. Publication bias complicates the interpretation of reviews and meta-analyses. In most cases overviews will be biased towards an overestimation of the effects of a treatment.

Contents of chapters 2-14

Chapter 2 discusses the methodology of searching the literature. It will be shown that a simple computer search is not sufficient to identify all relevant studies. Even articles which should readily be identified in computer searches, because they are published in indexed journals, will sometimes not be detected because of typing errors or misclassification. The number of relevant articles identified also depends heavily on the subject at hand. This may especially be the case for several food supplements.

Chapter 3 describes the results of a survey among general practitioners in the Netherlands. They were asked to what extent they believe in the efficacy of many alternative procedures. High scores were especially found for manual therapy, yoga, acupuncture, hot bath therapy and homoeopathy. Other procedures, such as many food supplements, were considered less useful.

In the chapters 4-12 the reviews of the efficacy of many food supplements for certain indications will be presented. It concerns vitamin C for the common cold (chapter 4), vitamin E for intermittent claudication and angina pectoris (chapter 5), vitamins B3 and B6 in mental functioning (chapter 6), and vitamin B6 for premenstrual syndrome (chapter 7), evening primrose oil for various indications (chapter 8), garlic for cardiovascular risk indicators (chapters 9 & 10) and ginkgo biloba for intermittent claudication and cerebral insufficiency (chapter 11). Also, in chapter 12, a review on the evidence of the efficacy of homoeopathy for various indications is presented. In these reviews, the methodological quality of the included trials will be emphasized. In general, the evidence is disappointing. Only for garlic, ginkgo and homoeopathy reasonable positive evidence exists. However, also for these three interventions, additional evidence from very well performed trials would be needed before the efficacy may be considered proven.

Chapter 13 discusses the problem of publication bias. Increasing numbers of articles dealing with this problem have appeared in the literature since 1980. Examples will be presented of surveys among researchers for unpublished trials and of comparisons of the

results of published and unpublished trials, showing that publication bias plays an important role in evaluating the scientific evidence.

Chapter 14 contains a general discussion of the most important problems and ideas which appeared while reviewing the literature of food supplements and after publishing our findings. Firstly, some limitations of criteria-based meta-analyses will be discussed. In the second place, we had the opportunity to visit many researchers and pharmaceutical firms, and thus we could find direct evidence that publication bias really exists for topics which are reviewed in this dissertation. Thirdly, we stress the importance of knowledge about the (postulated) mechanisms of action.

This dissertation ends with English and Dutch summaries (followed by a curriculum vitae and some acknowledgments). For modern reviews of the literature similar standards must be applied as for reports of controlled trials or other scientific research. This means a careful description and motivation of all relevant steps. It begins with the inclusion criteria and search strategies, followed by an assessment of the methodological quality of included trials, and ends with a presentation of the results with a discussion of the limitations of the review, and the potential influence of publication bias and methodological shortcomings.

References

1. Dorant E, Brandt van den PA, Hamstra AM, Feenstra MH, Bausch-Goldbohm RA. Gebruik van voedingssupplementen in Nederland. *Ned Tijdschr Geneesk* 1991;135:68-73.
2. Knipschild P. Ginseng: pep of nep? Een overzicht van experimenten bij ouderen met stoornissen van de vitaliteit. *Pharm Weekbl* 1988;123:4-11.
3. Knipschild P. Looking for gall bladder disease in the patient's iris. *Br Med J* 1988;297:1578-81.
4. Knipschild P. Changing belief in iridology after an empirical study. *Br Med J* 1989;299:491-2.
5. Riet ter G, Kleijnen J, Knipschild P. De effectiviteit van acupunctuur. *Huisarts Wet* 1989;32:(11 articles).
6. Riet ter G, Kleijnen J, Knipschild P. Acupuncture and chronic pain. A criteria-based meta-analysis. *J Clin Epidemiol* 1990;43:1191-9.
7. Riet ter G, Kleijnen J, Knipschild P. A meta-analysis of studies into the effect of acupuncture on addiction. *Br J Gen Pract* 1990;40:379-82.
8. Kleijnen J, Riet ter G, Knipschild P. Acupuncture and asthma. A review of controlled trials. *Thorax*: in press.
9. Knipschild P. Valkuilen van ongecontroleerd geneesmiddelonderzoek. In: Mochel H. e.a. *Chelatietherapie; een omstreden behandeling bij hart- en vaatziekten*. Kampen: Kok, 1988:64-79.
10. Riet ter G, Kleijnen J, Knipschild P. Over klinisch effectiviteitsonderzoek en placebo-effecten. *Integraal* 1988/89;(3):94-101.
11. Koes BW, Assendelft WJJ, Heijden van der GJMG, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *Br Med J*: in press.
12. Assendelft WJJ, Koes BW, Heijden van der GJMG, Bouter LM. The efficacy of chiropractic for back pain - blinded review of the relevant randomized clinical trials. *Journal of Manipulative and Physiological Therapeutics*: in press.
13. Kessels AGH, Cardynaals RLLM, Borger RLL, Knipschild PG et al. The effectiveness of the hair restorer "Dabao" in males with alopecia androgenetica. A clinical experiment. *J Clin Epidemiol* 1991;44:439-47.
14. Vet HCW de. The puzzling role of vitamin A in cancer prevention (review). *Anticancer Res* 1989;9:145-52.
15. Vet HCW de, Knipschild PG, Willebrand D, Schouten HJA, Sturmans F. The effect of beta-carotene on the regression and progression of cervical dysplasia: a clinical experiment. *J Clin Epidemiol* 1991;44:273-83.

CHAPTER 2

SEARCHING THE LITERATURE WITH ON-LINE COMPUTER DATABASES*

Jos Kleijnen, Paul Knipschild

* Submitted for publication

Summary

Objective - To assess the comprehensiveness of MEDLINE and EMBASE computer searches for controlled trials.

Design - Comparison of articles found after an exhaustive search of the literature with the yield of a MEDLINE or EMBASE search. This was performed for controlled clinical trials on the efficacy of three interventions: homoeopathy, ascorbic acid for common cold, and ginkgo biloba for intermittent claudication and cerebral insufficiency. The number of controlled trials found by exhaustive search of the literature was 107, 61 and 45 respectively.

Main outcome measures - The proportion of articles identified with MEDLINE and EMBASE respectively, and of articles found after stepwise checking of citations in MEDLINE and EMBASE articles. A comparison was made for all trials, and separately for the trials with higher methodological quality.

Results - For homoeopathy, ascorbic acid and ginkgo the proportion of all trials found by MEDLINE was 17%, 36% and 31% respectively and for EMBASE 13%, 25% and 58% respectively. After checking of the references in the MEDLINE articles 44%, 79% and 76% of all trials were identified. After checking of the references in the EMBASE articles 42%, 72% and 93% of all trials were identified. About 20% of the articles was not correctly indexed. Of the best trials 68%, 91% and 83% could be found with MEDLINE and 55%, 82% and 92% of the best trials were identified through EMBASE.

Conclusions - For the topics mentioned, MEDLINE and EMBASE searches are sufficient to get an impression of the evidence from controlled trials, but only if references in the articles are followed for further evidence. If one wants to get a more complete picture, additional search strategies make sense. Of course, this picture may be different for other topics.

Introduction

Nowadays, one of the first steps for the identification of relevant publications on a certain medical subject is searching of computerized data bases. More than 40 medical databases are currently operational, well known examples are MEDLINE from Index Medicus or EMBASE from Excerpta Medica.¹ Further checking of the references in identified articles often leads to a fair amount of extra information. The primary goal of this article is not to establish how correctly articles are indexed. The question is whether computerized searches for the retrieval of articles are a sufficient search method for reviewing the literature, or that additional efforts are necessary. Furthermore, if only part of the relevant articles is identified, it would be interesting to find out whether there are differences between the identified articles and other publications from non-indexed sources. Especially differences in the quality and in the outcome of trials would be interesting.

The comprehensiveness of computer searches will of course depend on the subject at hand. We chose for controlled clinical trials on the efficacy of three interventions: homoeopathy for various indications, ascorbic acid (vitamin C) for the prevention and treatment of common cold, and ginkgo biloba (leaves from the maidenhair tree) for intermittent claudication and cerebral insufficiency. We compared the number of articles found in a MEDLINE on-line search (1966-1991) and an EMBASE on-line search (1974-1991) with comprehensive review articles on these subjects.^{2,3,4} For these reviews exhaustive search strategies have been used, and each of these reviews exceeds by far the largest number of controlled trials cited in other reviews.

The effectiveness of MEDLINE searches

There have been several attempts to compare the result from MEDLINE searches with some other "gold standard" such as a registry of controlled trials, or comprehensive review articles. For instance, Dickersin et al sought for reports of randomized clinical trials of prevention and treatment of neonatal hyperbilirubinemia and prevention of intraventricular hemorrhage for the years 1966-1983 in MEDLINE files and in a register of controlled trials in perinatal medicine (National Perinatal Epidemiology Unit, Oxford, England). For trials of neonatal hyperbilirubinemia MEDLINE found only 42% of the 92 trials in the register. For trials of intraventricular hemorrhage, MEDLINE listed more trials than the 30 in the register.⁵

Hofmans used a series of 97 reports of controlled trials on the efficacy of acupuncture as a gold standard. These consisted of 1 unpublished study, 1 report, 8 abstracts, 4 publications in a series and 83 journal articles. Of these 83 articles, 67 were published in journals indexed in Index Medicus. A MEDLINE search yielded 46 articles (47% of the total). These included 18 (75%) of 24 studies considered to be of higher methodological quality.⁶

Methods

In the review articles which are used as a gold standard, exhaustive search strategies have been used. These include:

MEDLINE online search (1966-1991); EMBASE online search (1974-1991); Index Medicus manual search (only for vitamin C) (1940-1965); checking of Current Contents; checking references extensively in (review) articles on clinical research and in textbooks; checking the proceedings of conferences (only homoeopathy); checking the contents of several journals (only homoeopathy); personal communication with researchers; writing to and visiting major pharmaceutical companies; visiting several specialized libraries

(only vitamin C and homoeopathy); and telling people that you are looking for certain studies. By doing so we found 107 controlled trials on the efficacy of homoeopathy, 61 for vitamin C and common cold, and 45 for ginkgo in the treatment of intermittent claudication and cerebral insufficiency.

Keywords for the MEDLINE searches used for comparison with the above strategy were HOMEOPATHY, ASCORBIC ACID and RESPIRATORY TRACT INFECTIONS, GINKGO. These have been combined with the following set of keywords used for the identification of controlled trials: CLINICAL TRIALS, SINGLE-BLIND METHOD, EVALUATION STUDIES, FOLLOW-UP STUDIES, PROSPECTIVE STUDIES, COMPARATIVE STUDIES, DOUBLE BLIND METHOD. This quite elaborate computer search method was performed to make sure that all relevant studies could be detected. We thought it reasonable to assume that for indexing a controlled trial, both the intervention and the fact that it concerned some kind of comparison with another treatment or placebo, was indexed. Using this search method, we believed to compromise reasonably between comprehensiveness and efficiency.

Keywords for the EMBASE searches used for comparison with the above strategy were HOMEOPATHY, ASCORBIC ACID and RESPIRATORY TRACT INFLAMMATION, GINKGO and GINKGO BILOBA and GINKGO BILOBA EXTRACT. These have been combined with the following set of keywords used for the identification of controlled trials: CLINICAL TRIAL, FOLLOW UP, PROSPECTIVE STUDY, DOUBLE BLIND PROCEDURE, CROSS-OVER PROCEDURE and RANDOMIZATION. Only a limited number of trials was found using these keywords (5 trials of homoeopathy, 4 trials of ascorbic acid and 14 trials of ginkgo). When checking the list of indexed journals, it turned out that many other trials should have been identified. Subsequently, these trials were checked by using the first author's name. It appeared that the main keywords used in EMBASE for the identification of clinical trials are PLACEDBO and MAJOR CLINICAL STUDY. We did a new EMBASE search using these 2 keywords.

Results

As an illustration of how the total search is performed in practice, we present the full MEDLINE on-line search over the period January 1st 1966 - May 12th 1991:

1.00	570	CT=HOMEOPATHY
2.00	12666	CT=ASCORBIC ACID
3.00	168	FT=GINKGO
4.00	89938	FIND CT D RESPIRATORY TRACT INFECTIONS
5.00	78114	FIND CT D CLINICAL TRIALS
6.00	295	CT=SINGLE-BLIND METHOD
7.00	64134	CT=EVALUATION STUDIES
8.00	105762	CT=FOLLOW-UP STUDIES
9.00	36016	CT=PROSPECTIVE STUDIES
10.00	475087	CT=COMPARATIVE STUDIES
11.00	26740	CT=DOUBLE BLIND METHOD
12.00	691159	FIND 5 TO 11
13.00	259	FIND 2 AND 4
14.00	52	FIND 1 AND 12
15.00	81	FIND 12 AND 13
16.00	46	FIND 3 AND 12

CT means that the term is used as a keyword (MeSH term, Medical Subject Heading). FT (free text) means that the term can be located anywhere in the text. 5 TO 11 means: 5 or 6 or 7 or 8 or 9 or 10 or 11. D means that all relevant subsets of a term will be searched. These are for CLINICAL TRIALS: MULTICENTER STUDIES and RANDOMIZED CONTROLLED TRIALS, and for RESPIRATORY TRACT INFECTIONS many terms, including COMMON COLD.

Thus we found 52 references for homoeopathy, 81 for vitamin C and respiratory tract infections, and 46 for ginkgo which might be controlled trials.

The EMBASE search was performed using exactly the same scheme, but of course with the keywords (which we already described) appropriate for EMBASE.

Homoeopathy

MEDLINE

Of the 52 identified references, 18 were indeed controlled trials of the efficacy of homoeopathy. Compared with the exhaustive search (107 trials in 96 publications) this is 17% of all controlled trials. When checking the list of indexed journals, it turned out that 5 trials which should have been identified, were not found.^{7,8,9,10,11} Subsequently, these trials were checked by using the first author's name. For one trial, no term had been used to indicate that it was a controlled trial,⁹ and for the other four the keyword PLANT EXTRACTS had been used instead of HOMEOPATHY.

Next came stepwise checking of references, starting with the 18 identified studies. This yielded 12 additional studies, including all five indexed trials which had not been found initially (table). From these 12 studies, 8 were found in the first check of references, another 2 studies were found as a reference in the 8 studies of the first check, and 2 studies were found in the 2 of the second check. There were also references to two books on homoeopathy in which another 16 studies could be found. Thus, all in all, the MEDLINE search and consequent checking of references yielded 46 publications reporting 47 controlled trials. This is 44% of all trials and 48% of all publications.

From the 22 studies that turned out to be of higher methodological quality (15 reported positive outcomes and 7 reported negative outcomes), one third (5 positive and 2 negative) was missed.

EMBASE

Of the 74 identified references, 14 were indeed controlled trials of the efficacy of homoeopathy. Compared with our exhaustive review (107 trials in 96 publications) this is 13% of all controlled trials. When checking the list of indexed journals, it turned out that 10 publications containing 10 trials which should have been identified, had not been found.¹²⁻²¹ Subsequently, these trials were checked by using the first author's name. In 8 publications there was no indexed term of homoeopathy, in 1 there was no indexed term indicating that it might concern a controlled trial, and one publication was a letter to the editor, which was probably not indexed.

Next came stepwise checking of references (table), starting with the 14 identified studies. This yielded 17 additional studies (including 4 indexed publications (4 trials) not found initially).^{12,15,18,19} There were also references to two books on homoeopathy in which another 13 studies could be found. Thus, all in all, the EMBASE search and consequent checking of references yielded 44 publications reporting 45 controlled trials. This is 42% of all trials and 46% of all publications.

From the 22 studies which had the highest methodological quality (15 reported a positive outcome and 7 reported negative outcomes), 10 studies (6 positive and 4 negative) were missed.

Table. Checking references in identified articles

	MEDLINE	EMBASE
Homoeopathy		
Initial search	18	14
First check	8	10
Second check	2	3
Third check	2	4
Books	16	13
Total	46 (44%)	44 (42%)
Ascorbic acid		
Initial search	22	15
First check	15	16
Second check	9	11
Third check	1	2
Total	47 (77%)	44 (72%)
Ginkgo biloba		
Initial search	14	26
First check	17	11
Second check	3	5
Third check	0	0
Total	34 (76%)	42 (93%)

*Ascorbic acid and common cold***MEDLINE**

Of the 81 identified references, 22 were indeed controlled trials of the efficacy of vitamin C for common cold. Compared with our exhaustive review (61 trials in 60 publications) this is 36% of all controlled trials. When checking the list of indexed journals, it turned out that 5 publications containing 6 trials which should have been identified, were not found.²²⁻²⁶ Subsequently, these trials were checked by using the first author's name. In all 5 publications, no term had been used to indicate that it was a controlled trial. Next came stepwise checking of references, starting with the 22 identified studies (table). This yielded 25 additional studies (including two indexed publications (2 trials) not found initially).^{22,23} Thus, all in all, the MEDLINE search and consequent checking of references yielded 47 publications reporting 47 controlled trials. This is 77% of all trials and 78% of all publications.

From the 11 studies which had the highest methodological quality (7 reported a small positive outcome in the treatment of colds, and 4 reported negative outcomes), only 1 positive study was missed.

EMBASE

Of the 57 identified references, 15 were indeed controlled trials of the efficacy of vitamin C for common cold. Compared with our exhaustive review (61 trials in 60 publications) this is 25% of all controlled trials. When checking the list of indexed journals, it turned out that 4 publications containing 5 trials which should have been identified, were not found.^{24,25,27,28} Subsequently, these trials were checked by using the first author's name. The first publication (from 1974) was not found at all.²⁷ The second publication had only been indexed with the term drug efficacy to indicate that it concerned a controlled trial.²⁵ The

other 2 publications had not been indexed with a term for common cold or respiratory tract inflammation.^{24,28}

Next came stepwise checking of references, starting with the 15 identified studies (table). This yielded 29 additional studies (including 2 indexed publications (2 trials) not found initially)^{27,28}. Thus, all in all, the EMBASE search and consequent checking of references yielded 44 publications reporting 44 controlled trials. This is 72% of all trials and 73% of all publications.

From the 11 studies which had the highest methodological quality (7 reported a small positive outcome in the treatment of colds, and 4 reported negative outcomes), only 2 studies (one positive and 1 negative) were missed. The EMBASE search did not yield any new, previously unknown studies, although there were a couple of duplicate publications that we didn't know of.

Ginkgo biloba for intermittent claudication and cerebral insufficiency.

MEDLINE

Of the 46 identified references (in May 1991), 14 were indeed controlled trials of the efficacy of ginkgo. Compared with our exhaustive search (45 trials in 45 publications) this is 32% of all controlled trials. When checking the list of indexed journals, it turned out that 4 publications which should have been identified, were actually not found.^{29,32} Subsequently, these were checked by using the first author's name. In 2 of these publications, no term had been used to indicate that it concerned ginkgo, in one publication the word had been spelled as *gingko*,²⁹ and in another as *gingkco*.³⁰

Next came stepwise checking of references, starting with the 14 identified studies (table). This yielded 20 additional studies (including 3 indexed publications not found initially).^{29,31,32} Thus, all in all, the MEDLINE search and consequent checking of references yielded 34 controlled trials. This is 76% of all publications. From the 12 studies which were of higher methodological quality (all reported positive outcomes), only 2 were missed.

EMBASE

Of the 89 identified references, 26 were indeed controlled trials of the efficacy of ginkgo. Compared with our exhaustive review (45 trials in 45 publications) this is 58% of all controlled trials. When checking the list of indexed journals, it turned out that 7 publications containing 7 trials which should have been identified, were not found.^{31,33-38} Subsequently, these trials were checked by using the first author's name. One of the studies was indexed as *gingho bilola*, four were not found at all, in one study there was no indication that it concerned ginkgo and, finally, in one study the only indication that it concerned a controlled trial was the term DRUG COMPARISON.

Next came stepwise checking of references, starting with the 26 identified studies (table). This yielded 16 additional studies (including 6 indexed publications (6 trials) not found initially).^{31,33,34,36-38} Thus, all in all, the EMBASE search and consequent checking of references yielded 42 publications reporting 42 controlled trials. This is 93% of all trials and of all publications. From the 12 studies which had the highest methodological quality (all reported a positive outcome), only 1 study was missed.

Discussion

Estimation of the "yield" of computer searches of course depends on the completeness of the exhaustive searches. For our three comparisons, and for the comparisons made by Hofmans (the exhaustive search was also a review written by us),³⁹ the reviews were the

most comprehensive ever published. For all these subject matters collection of publications took place over a period of several years, and after the reviews had been published, no omissions were noticed by other persons.

For a search strategy such as used to identify articles, there is a trade-off between the completeness (the proportion for the indexed articles identified) and the noise (the proportion of useless articles found). In most cases, the noise must not be too much (maximally a few hundreds?), because it is inefficient to screen too many useless articles. The initial yield in our MEDLINE search was 125 useless articles and 54 articles we wanted to identify, for the EMBASE search the corresponding numbers were 220 and 55. Consequent checking all citations in the identified articles and of all subsequent citations yielded another 74 articles for MEDLINE and another 75 articles for EMBASE. Clearly checking references is very useful.

Fourteen of the 68 MEDLINE articles (21%) appeared to be not correctly indexed (assuming that our search strategy was reasonably optimal). In two cases the authors spelled the intervention wrong (ginkgo), and in 6 of the other 12 missing articles there was no indication that it concerned a controlled trial. For homoeopathy it could be argued that we should also have used the keyword PLANT EXTRACTS (sometimes the difference between homoeopathy and phytotherapy is hard to see); by not doing so we missed 4 articles. Twenty-one of the 76 EMBASE articles (28%) appeared to be not correctly indexed (assuming that our search strategy was reasonably optimal). The reasons were similar as with MEDLINE. It was interesting to see that the keywords which in MEDLINE are used to identify controlled trials (CLINICAL TRIALS, DOUBLE BLIND METHOD, COMPARATIVE STUDIES etc.) do exist in EMBASE (be it slightly modified), but surprisingly do not yield additional articles to the ones found with the keywords PLACEBO and MAJOR CLINICAL STUDY.

We also found several review articles from other authors. For instance, the most comprehensive review that we found on vitamin C and common cold, was in a book by Pauling.⁴⁰ He cited 32 controlled trials, other reviewers of this subject cited fewer studies. However, checking citations in the articles identified in a MEDLINE search would have yielded 16 extra controlled trials (the book which contained the review was published in 1986; the most recent trial was published in 1982). For most reviews of homoeopathy and ginkgo the same amount of underreporting was found. The inevitable conclusion must be that many reviewers either do not check references adequately, or that they present only a (biased) selection of all the evidence.

The methodological quality of the studies had been assessed in our reviews. Most of the best studies of all 3 topics could be identified in the MEDLINE and EMBASE searches, only if these were followed by checking of citations: for homoeopathy 68% and 55% respectively, for ascorbic acid 91% and 82%, and for ginkgo 83% and 92%. If one wants to get a more complete picture, additional search methods should be used. Of course, this picture may be different for other topics.

References

1. Vogten AJM. Medical information - online. A critical review on computer access to medical information. *Neth J Med* 1988;32:34-49.
2. Kleijnen J, Knipschild P, Riet ter G. Clinical trials of homoeopathy. *Br Med J* 1991;302:316-23.
3. Kleijnen J, Riet ter G, Knipschild P. Vitamine C en verkoudheid. Overzicht van een megadosis literatuur. *Ned Tijdschr Geneesk* 1989;133:1532-5.
4. Kleijnen J, Knipschild P. Ginkgo biloba for intermittent claudication and cerebral insufficiency. In: Food supplements and their efficacy [Dissertation]. Maastricht: University of Limburg, 1991.
5. Dickersin K, Hewitt P, Mutch L, Chalmers I, Chalmers TC. Perusing the literature: Comparison of MEDLINE searching with a perinatal trials database. *Controlled Clinical Trials* 1985;6:306-17.

6. Hofmans EA. De opbrengst van een MEDLINE-search. De toegankelijkheid van onderzoek naar de effectiviteit van acupunctuur II. Huisarts en Wetenschap 1990;33(3):103-6.
7. Rahlfs WV, Mössinger P. Zur Behandlung des Colon irritabile. Arzneimittelforschung 1976;26:2230-4.
8. Rahlfs VW, Mössinger P. Asa foetida bei Colon irritabile. Doppelblindversuch. Dtsch Med Wochenschr 1979;104:140.
9. Gassinger CA, Wünstel G, Netter P. Klinische Prüfung zum Nachweis der therapeutischen Wirksamkeit des homöopathischen Arzneimittels Eupatorium perfoliatum D 2 (Wasserhanf composite) bei der Diagnose "Grippaler Infekt". Arzneimittelforschung 1981;31:732-6.
10. Wiesenauer M, Häussler S, Gaus W. Pollinosis-Therapie mit Galphimia glauca. Fortschr Med 1983;101:811-4.
11. Kubista E, Müller G, Spona J. Behandlung der Mastopathie mit cyclische Mastodynie: klinische Ergebnisse und Hormonprofile. Gynäkologische Rundschau 1986;26:65-79.
12. Wiesenauer M, Häussler S, Gaus W. Pollinosis-Therapie mit Galphimia glauca. Fortschr Med 1983;101:811-4.
13. Maiwald L, Weinfurter T, Mau J, Connert WD. Therapie des grippalen Infekts mit einem homöopathischen Kombinationspräparat im Vergleich zu Acetylsalicylsäure. Kontrollierte, randomisierte Einfachblindstudie. Arzneimittelforschung 1988;38:578-82.
14. Wiesenauer M, Gaus W, Bohnacker U, Häussler S. Wirksamkeitsprüfung von homöopathische Kombinationspräparaten bei Sinusitis. Ergebnisse einer randomisierten Doppelblindstudie unter Praxisbedingungen. Arzneimittelforschung 1989;39:620-5.
15. Chevrel JP, Saglier J, Destable MD. Reprise du transit intestinal en chirurgie digestive. Action homéopathique de l'Opium. Presse Med 1984;13:833.
16. Wiesenauer M, Gaus W. Orthostatische Dysregulation. Kontrollierter Wirkungsvergleich zwischen Etilefrin 5 mg und dem homöopathischen Arzneimittel Haplopappus D2. Zeitschrift für Allgemeinmedizin 1987;63:18-23.
17. Lewith G, Brown PK, Tyrell DAJ. Controlled study of the effects of a homoeopathic dilution of influenza vaccine on antibody titres in man. Complementary Medical Research 1989;3:22-4.
18. Rahlfs WV, Mössinger P. Zur Behandlung des Colon irritabile. Arzneimittelforschung 1976;26:2230-4.
19. Rahlfs VW, Mössinger P. Asa foetida bei Colon irritabile. Doppelblindversuch. Dtsch Med Wochenschr 1979;104:140.
20. Fisher P, Greenwood A, Huskisson EC, Turner P, Belon P. Effect of homoeopathic treatment on fibrositis (primary fibromyalgia). Br Med J 1989;299:365-6.
21. Leaman AM, Gorman D. Cantharis in the early treatment of minor burns. Arch Emerg Med 1989;6:259-61.
22. Walker GH, Bynoe ML, Tyrrell DAJ. Trial of ascorbic acid in the prevention of colds. Br Med J 1967;i:603-6.
23. Charleston SS, Clegg KM. Ascorbic acid and the common cold. Lancet 1972;i:1401-2.
24. Gormly PJ. Megadose of ascorbic acid in an Antarctic expedition. Br J Nutr 1977;37:269-77.
25. Ludvigsson J, Hansson LO, Tibbling G. Vitamin C as a preventive medicine against colds in children. Scand J Infect Dis 1977;9:91-8.
26. Martin NG, Carr AB, Oakeshott JG, Clark P. Co-twin control studies: vitamin C and the common cold. Prog Clin Biol Res 1982;103a:365-73.
27. Coulehan JL, Reisinger KS, Rogers KD, Bradley DW. Vitamin C prophylaxis in a boarding school. N Engl J Med 1974;290:6-10.
28. Anderson TW, Beaton GH, Corey PN, Spero L. Winter illness and vitamin C: the effect of a relatively low dose. Can Med Assoc J 1975;112:823-6.
29. Halama P, Bartsch G, Meng G. Hirnleistungsstörungen vaskulärer Genese. Randomisierte Doppelblindstudie zur Wirksamkeit von Ginkgo-biloba-Extrakt. Fortschr Med 1988;106:408-13.
30. Thomson GJL, Vohra RK, Carr MH, Walker MG. A clinical trial of Ginkgo biloba extract in patients with intermittent claudication. Int Angiol 1990;9:75-8.
31. Moreau P. Un nouveau stimulant circulatoire cérébral. Presse Med 1975;4:2401-2.
32. Eckmann F, Schlag H. Kontrollierte Doppelblind-Studie zum Wirksamkeitsnachweis von Tebonin forte bei Patienten mit zerebrovaskulärer Insuffizienz. Fortschr Med 1982;100:1474-8.
33. Ambrosi C, Bourde C. Nouveauté thérapeutique médicale dans les artériopathies des membres inférieurs: Tanakan. Essai clinique et étude par les cristaux liquides. Gazette Médicale de France. 1975;82:628-32.

34. Bastide G, Montsarrat M. Artérite des membres inférieurs. Intérêt du traitement médical après intervention chirurgicale. Analyse factorielle. *Gazette Médicale de France* 1978;85:4523-6.
35. 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.
36. Augustin P. Le Tanakan en gériatrie. Etude clinique et psychométrique chez 189 malades d'hospice. *Psychologie Médicale* 1976;8:123-30.
37. 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.
38. 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 Médical* 1977;13:1197-9.
39. Riet ter G, Kleijnen J, Knipschild P. De effectiviteit van acupunctuur. *Huisarts Wet* 1989;32:(11 articles).
40. Pauling L. How to live longer and feel better. New York: W.H. Freeman and Company, 1986.

RESEARCH NOTE

BELIEF IN THE EFFICACY OF ALTERNATIVE MEDICINE AMONG GENERAL PRACTITIONERS IN THE NETHERLANDS

PAUL KNIPSCHILD, JOS KLEIJNEN and GERBEN TER RIET

Department of Epidemiology/Health Care Research, University of Limburg, P.O. Box 616, 6200 MD
Maastricht, the Netherlands

Abstract—A survey among 293 GPs in the Netherlands showed that many believe in the efficacy of common alternative procedures. High scores were especially found for manual therapy, yoga, acupuncture, hot bath therapy and homoeopathy. Other procedures, such as iridology, faith healing and many food supplements, were considered less useful.

Key words—alternative medicine, efficacy, general practice, questionnaire

INTRODUCTION

Alternative medicine is quite popular among GPs in the United Kingdom [1-3]. An earlier study in the Netherlands showed that almost all GPs refer certain patients to alternative practitioners, and that half of them practice alternative medicine themselves (mostly homoeopathy) [4]. This policy may partly be caused by patients' pressure, financial motives or reluctance to prescribe pure placebos. However, this report shows that many GPs really believe in the efficacy of certain alternative procedures.

DOCTORS AND METHODS

From the register of GPs in the Netherlands a random sample of 400 doctors was drawn. In the summer of 1989 they received a postal questionnaire to state their belief in the efficacy of many alternative procedures [5]. It was stressed that our interest was in specific effects, beyond placebo effects, and that all procedures should be considered under optimum circumstances (complete compliance, expert practitioner). They could give their opinion on a scale from 0 to 10. The familiar report marks should not indicate perceived quantitative effects, but the degree of belief in any efficacy. Small but almost certain effects should be rated much higher than large but seriously doubted effects (say, 9 vs 2).

RESULTS

Five addresses turned out to be wrong. After two reminders the response was 74% or 293 doctors. Table 1 shows their scores (mean and distribution). In the text we only mention the percentage of positive scores (6-10); the width of its 95% confidence interval is approx. 10%. The scores of various alternative procedures are reported under seven headings.

Acupuncture

Especially for patients with chronic pain, acupuncture is considered efficacious by half of the doctors. Some 25% believe acupuncture to be a useful treatment for asthma or smoking addiction. Neural therapy is positively scored by 15%.

Homoeopathy

About 45% thinks that homoeopathic remedies are efficacious in the treatment of upper respiratory tract infections or hay fever. For chronic joint problems this is 30%. Anthroposophical medicine makes sense to 22%.

Spiritual techniques

Compared to acupuncture and homoeopathy doctors give less credit to faith healing (7%) or healing (17%). Nevertheless, 30% report to believe that hypnosis can help to quit smoking. For patients with psychological problems yoga appeals to 62%, and more conventional psychotherapy to 83%.

Manual techniques

Manual therapy in the Netherlands is generally not considered alternative medicine anymore. No less than 80% believes it to be efficacious in the treatment of patients with chronic neck or back problems, compared to 71% for conventional physiotherapy. Hot bath (Kurort) therapy is positively rated by 46% for rheumatic patients. With 14% reflexology (foot massage) scores much lower.

Natural remedies

Among the many available products we inquired after evening primrose oil (premenstrual syndrome, atopic eczema), garlic (cardiovascular disease) and ginseng (decreased vitality). Less than 10% considers them efficacious. The 'Moerman' diet is well-known

Table 1. Report marks of Dutch GPs for their belief in the efficacy of alternative medicine

Procedure	Indication	Mean score	Distribution (%)					No.*
			0-1	2-3	4-5	6-7	8-10	
Acupuncture	Chronic pain	5.0	12	12	26	36	14	291
Acupuncture	Asthma	3.1	35	19	23	17	5	279
Acupuncture	Smoking	3.7	26	22	24	20	8	288
Neural therapy	Pain	2.3	51	15	19	11	4	268
Homoeopathy	Upper respiratory infection	4.7	20	10	24	32	15	287
Homoeopathy	Hay fever	4.5	20	11	25	32	11	290
Homoeopathy	Chronic joint disease	3.7	26	19	25	22	7	287
Anthroposophy	General	3.2	35	17	25	18	4	271
Faith healing	General	1.5	65	17	11	4	3	282
Healing	General	2.6	43	18	23	14	3	287
Hypnosis	Smoking	3.6	28	18	24	26	4	275
Meditation, yoga	Psychological problems	5.6	7	11	21	43	19	290
Psychotherapy	Psychological problems	6.7	2	2	12	50	33	291
Physiotherapy	Chronic neck/back problems	6.3	3	4	21	43	28	290
Manual therapy	Chronic neck/back problems	6.4	4	4	12	47	33	289
Hot bath (Kurort)	Rheumatism	4.6	14	17	23	38	8	278
Reflexology	General	2.4	49	16	21	11	3	276
Evening primrose	Premenstrual syndrome	1.5	64	16	16	3	1	259
Evening primrose	Atopic eczema	1.6	62	15	17	4	2	258
Garlic	Cardiovascular disease	1.5	63	19	10	8	0	286
Ginseng	Decreased vitality	1.7	60	18	13	9	0	280
Moerman diet	Cancer	1.9	56	18	14	11	1	289
Calcium (megadose)	Cancer prevention	0.9	76	14	8	1	0	271
Beta-carotene	Cancer prevention	1.3	72	12	10	6	1	267
Vitamin C	Cancer prevention	1.6	62	17	12	7	3	280
Vitamin E	Atherosclerosis	1.3	68	17	10	5	0	275
Vitamin comb	General prevention	2.2	51	20	16	12	2	281
Vitamin C	Common cold	2.5	47	16	17	15	5	283
Vitamin B6	Premenstrual syndrome	3.9	28	15	22	25	11	282
Astrology	Diagnosis	0.8	80	12	4	3	1	278
Iridology	Diagnosis	1.4	68	14	8	7	2	285
Ozon therapy	Cardiovascular disease	0.8	77	16	5	2	0	265
Chelation therapy	Cardiovascular disease	1.8	59	17	16	7	1	281
Enzyme therapy	General	1.3	68	17	10	4	0	277
Cell therapy	General	1.0	75	13	10	1	0	277
Thymus extract	General	0.8	80	12	8	1	0	260

*For separate items the response is lower than 293 out of 395.

in the Netherlands as an alternative therapy for cancer patients. However, only 12% of the GPs believes it to have any efficacy.

Orthomolecular medicine

The same negative opinion is usually expressed for megadoses of calcium, beta-carotene or vitamin C (cancer prevention), vitamin E (atherosclerosis) or combinations of vitamins (general prevention). Vitamin C in megadose is considered efficacious by 20% for common cold, whereas vitamin B6 in cases of the premenstrual syndrome is positively scored by 35%.

Miscellaneous

Many other procedures are generally believed to make no sense. Among the diagnostic procedures less than 10% considers astrology or iridology to be proper medical practice. The same goes for ozon therapy and chelation therapy (cardiovascular disease), and for enzyme therapy, cell therapy or injections with thymus extract.

DISCUSSION

Our study shows that many Dutch GPs believe in the efficacy of common alternative procedures. Of course, it is only a survey of opinions and no conclusions can be drawn about true efficacy.

We are surprised about the high amount of credit that is given to certain (to us) incomprehensible

practices such as acupuncture, homoeopathy and anthroposophy. It must be realized that even low percentages of positive scores represent a good many doctors. Especially if there is a large contrast between belief in and evidence of efficacy, the publication of literature reviews and empirical studies in scientific journals and the lay press makes sense.

Acknowledgements—This study was sponsored by the Netherlands Ministry of Welfare, Public Health and Cultural Affairs. Additional support was given by five firms of food supplements. The random sample of Dutch GPs was drawn by NIVEL.

REFERENCES

1. Reilly D. T. Young doctors' views on alternative medicine. *Br. Med. J.* 287, 337-339, 1983.
2. Wharton R. and Lewith G. Complementary medicine and the general practitioner. *Br. Med. J.* 292, 1498-1500, 1986.
3. Anderson E. and Anderson P. General practitioners and alternative medicine. *J. R. Coll. Gen. Pract.* 37, 52-55, 1987.
4. Visser J. and Peters L. Alternative medicine and general practitioners in the Netherlands: towards acceptance and integration? *Fam. Pract.* In press.
5. *The Handbook of Complementary Medicine* (Edited by Fulder S.). Hodder & Stoughton/Coronet Books, London, 1984.

CHAPTER 4

VITAMIN C AND THE COMMON COLD; REVIEW OF A MEGADOSE OF LITERATURE*

Jos Kleijnen, Gerben ter Riet, Paul Knipschild

* Published in Ned Tijdschr Geneesk 1989;124:418-423

Introduction

Nowadays, many people use extra vitamins, minerals and other products which may cause health improvement. There is a good deal of discussion about which benefits to expect. The question is whether the recommended dietary allowances, which are aimed at the prevention of deficiency symptoms, are also the dosages necessary for optimal health. One of the most extensively investigated claims is the influence of megadosages of vitamin C on the common cold. Despite many publications during the years past, discussions about this topic are still ongoing.

In this article we describe how we retrieved the publications on vitamin C and the common cold, and how we made a selection of those with the highest quality using a criteria list. The results of these studies were used for the assessment of the influence of vitamin C supplementation on the common cold.

Methods

Retrieval of the literature

Many review articles on vitamin C and the common cold only describe a limited number of trials.¹⁻⁷ There are, however, many other articles on vitamin C and the common cold which have been published, especially in journals and languages which are not readily accessible. We identified 61 trials by checking of the Index Medicus (1970-1987), and checking of references in all articles found.⁸⁻²² We listed all trials in the reference list. Probably, this is not a complete list of all performed trials. We checked this approach by means of a MEDLINE-computersearch (1963-1988). No additional articles were found using the keywords "ascorbic acid", "common cold", and "respiratory tract infections". Publications were included in this review if a group treated with vitamin C only was compared with a control group.

Selection of the publications

We wanted to answer the following questions:

- Is there an influence of vitamin C in the prevention of colds?
- Is there an influence of vitamin C in the treatment of colds?
- Which dosage is effective?
- In which people are any effects found?

We made a list of criteria of good methodology, because reliable results can only be obtained from trials with sound methodology. Publications were selected for a detailed review using the following criteria list. We think that randomization and blinding are most important:

- Randomized trials could earn 3 points if the method of randomization was described; if the randomization was not described, 2 points were given; and alternate allocations were given 1 point.
- Double-blind trials earned 3 points, and single-blind trials 1 point.
- Furthermore, 1 point each was given if blinding had been checked, if the placebo had been described, if there were at least 100 participants per group, if drop-outs had been described (and if there were no more than 25% drop-outs), if the intervention lasted at least 3 months (for prophylaxis) and if there was a therapeutical follow-up of at least one week. Finally, a point was given if both the endpoints had been described and there was a differentiation between subsequent episodes of symptoms.

In this way, maximally 12 points could be earned. There were 11 trials which earned at least 8 points, and these trials will be described in detail in this review. The trials of Wilson et al.,⁴⁴⁻⁴⁶ Tyrell et al.⁴³ and McLean Baird et al.⁴⁴ earned 7 points, as well as a

second trial of Ludvigsson et al., which was published in the same article as the trial presented in the table.⁶¹ The other trials scored 6 points or fewer. In order to give an example of the trials involved, we discuss aspects of the trial by Pitt and Costrini in some detail. Together with the trial of Coulehan et al.⁵³ this trial earned 10 points, which was the highest score.

In the double-blind trial of Pitt and Costrini 862 recruits were treated during an 8 week period. The success of the randomization procedure was checked by comparison of 7 characteristics which were measured at baseline. There were no relevant differences between the intervention and placebo groups. Vitamin C was administered in a dosage of 2.0g/day. The placebo tablets were formulated from citric acid and were indistinguishable in appearance and taste from the vitamin C tablets. The criteria necessary for a cold to be included as an episode were (1) the presence of either runny or stuffy nose, sore throat, or dry or productive cough, (2) at least two days of symptoms, and (3) at least three symptom-free days between episodes. Eight different symptoms were followed-up by means of weekly questionnaires. It was checked whether the pills had actually been taken and the recruits were asked to report side effects and any lapses in pill taking. On the final questionnaire, the recruits were asked to state if they knew which pill they had been taking and how they knew. Sixty-four recruits (34, vitamin C; 30, placebo) were removed from their platoons. An additional 123 recruits (64, vitamin C; 59, placebo) were excluded because they did not continue to take their pills and 1 recruit from the vitamin C group dropped out because of recurrent urticaria related to taking the tablets. Thus, unfortunately, there was no intention-to-treat analysis. There were no apparent differences for the incidence of colds, and the duration and severity of the symptoms. The authors concluded that vitamin C is not effective in the prevention of colds.

Results

The table shows that the treated groups vary from schoolchildren and soldiers to twins and employees of large companies. All 11 double-blind trials presented in the table, assessed the prophylactic as well as the therapeutic effect of vitamin C. The dosage of vitamin C which was used after 1960 is mostly 1 g/day or more. Before 1960, dosages up to 200mg/day were regular.

There are striking differences between the endpoints used: "colds", "symptoms of the nose or throat", "respiratory illness" and "winter illness". Even greater differences are encountered regarding the duration and the severity of the symptoms; sometimes it was not clear whether the endpoint concerned the duration or the severity of the cold. These differences complicate a definitive assessment of the results of these trials.

Only Bancalari et al. found a small positive effect on the incidence of colds.⁶⁰ Regarding the therapeutical effects of vitamin C we find both for the duration and the severity of the symptoms a small positive effect (approx. 10% or more) in seven trials, and no effect in 4 trials.

From these trials, it is not possible to make a reliable assessment of the optimal dosage needed. It appears that the results for dosages of 1 g/day or more and 200 mg/day are similar. Moreover, it is impossible to establish whether there are any subgroups who might have more or less benefit from vitamin C supplementation, because the differences between trials regarding the intervention and the effect measurement are too large.

authors, year, country	population	number of participants (dropouts)	duration in weeks
Anderson et al. 1972 Canada ³⁹	employees of several companies	407/411 (182)	9-17
Anderson et al. 1974 Canada ⁴⁷	staff of hospitals and several compa- nies	approx. 300 per group total 2349 (1171)	13
Coulehan et al. 1974 USA ⁴⁹	schoolchildren	321/320 (13/12)	14
Anderson et al. 1975 Canada ⁵¹	staff of hospitals, large companies and students	tablets 150, capsules 152/ placebo 146 (57,56/61)	15
Karlowski et al. 1975 USA ⁵⁴	employees Natl. Inst. Health	totally 311 (more dropouts in placebo group)	39
Coulehan et al. 1975 USA ^{53,55}	schoolchildren	428/428 combination of 2 trials (88)	15-18
Ludvigsson et al. 1977 Sweden ⁶¹	schoolchildren	304/311 (27)	13
Miller et al. 1977 USA ⁶²	twins	42 pairs (2 pairs)	22
Pitt & Costrini 1979 USA ⁶⁵	recruits	331/343 (99/89)	8
Carr et al. & Martin et al. 1981, 1982 Australia ⁶⁶⁻⁶⁸	twins	95 pairs (30/30)	14
Bancalari et al. 1984 Chile ⁶⁹	schoolchildren	32/30 no dropouts	12

"/" divides vitamin C group from placebo group

daily dosis of vitamin C	symptoms of disease		
	number of colds per person	severity	duration (days)
prophyl.; therap. 1 g; 4 g	1,1/1,2	vit. C-group less severe	4,6/5,4
prophyl.; therap. 1. 1 g; 4 g 2. 1 g; - 3. 2 g; - 4. - ; - 5. 250 mg; - 6. - ; - 7. - ; 4 g 8. - ; 8 g	no difference	no difference	no difference
1. 1 g 6-10 years 2. 2 g 10-15 years	1. 0.1/0.1 2. 0.1/0.1	vit. C-groups less severe	1. 5.0/5.7 2. 4.4/6.3
prophyl.; therap. 500 mg per week; 1500 mg on day 1 1000 mg on day 2-5	not mentioned	vit. C-groups less severe	1.2, 1.2/1.6
prophyl.; therap. 1. 3 g; 3 g 2. 3 g; - 3. - ; 3 g 4. - ; -	1. 1.3 2. 1.2 3. 1.3 4. 1.4	vit. C-group less severe	1. 5.9 2. 6.7 3. 6.5 4. 7.1
1 g	0.2/0.2	no difference	5.5/5.8
1 g placebo contains 10 mg vit. C	2.2/2.0	vit. C-group less severe	6.0/5.7
depending on body weight 0.5, 0.75 or 1.0 g (all groups multi-vitamins with additional 50 mg vit.C)	5.0/4.8	vit. C-group less severe	6.9/7.5
2 g	1.8/1.8	no difference	11.2/11.5
1 g (all groups multi-vitamins with additional 50 mg vit.C)	1.6/1.5	vit. C-group less severe	5.2/6.4
2 g	1.2/1.5	vit. C-group less severe	3.4/4.5

* divides vitamin C group from placebo group

Discussion

It appears that a great deal of research has been performed on the effects of vitamin C on colds. Of course, we can not exclude the possibility that some trials with "negative" results have remained unpublished.

A look at the populations studied during the years shows us that research on vitamin C and colds has been performed in a great variety of people. This would enable us to draw conclusions about subgroups that might benefit of vitamin C suppletion. However, not only the people participating in these experiments varied, but also many other aspects of the experiments. The same applies to other interesting questions, such as a dose-response effect.

Furthermore, the methodological quality of the earlier trials is disappointing. No trial performed before 1970 attained a high methodological score.

The outcome of the experiments might be summarized as follows: there is no prophylactic effect of vitamin C on the incidence of common colds, but as a therapy it might exert some effect. A positive effect of approx. 10% both on the duration and on the severity of the symptoms cannot be excluded.

The implications for general practice may depend on the perspective from which one looks at this evidence. The patient has to weigh the costs and effort of vitamin C supplementation against the small benefit which can be expected. From the doctor's point of view, there is not much to be gained, because colds are self-limiting. A gain of half a day is not much for a cold that lasts for five days. On the other hand, considering the high incidence of colds, a gain of half a day sick-leave might very well be worthwhile from a larger, nationwide, perspective.

A more comprehensive report of our findings can be found in an earlier published report (in Dutch).⁷³

This work was supported by a grant of the Dutch Ministry of Welfare, Public Health and Cultural Affairs.

References

1. Anonymous. Ascorbic acid and the common cold. *Nutr Rev* 1967;25:228-31.
2. Pauling L. Vitamin C and the common cold. San Francisco: Freeman, 1970.
3. Ketz HA. Über die prophylaktische Wirkung des Vitamin C und den Vitamin-C-Bedarf des gesunden Erwachsenen. *Zeitschrift für Medizin* 1970;25:414-7.
4. Dykes MHM, Meier P. Ascorbic acid and the common cold. Evaluation of its efficacy and toxicity. *JAMA* 1975;231:1073-9.
5. Chalmers TC. Effects of ascorbic acid on the common cold. An evaluation of the evidence. *Am J Med* 1975;58:532-6.
6. Wilson CWM. Colds, ascorbic acid metabolism, and vitamin C. *J Clin Pharmacol Ther* 1975;15:570-8.
7. Pauling L. How to live longer and feel better. New York: Freeman, 1986.
8. Ruskin SL. Calcium cevitamate in the treatment of acute rhinitis. *Ann Otol Rhinol Laryngol* 1938;47:502-11.
9. Peters M. Zur Frage der resistenzsteigernden Wirkung prophylaktischer Dauermedikation von Vitamin C. *Zeitschrift für Kinderheilkunde* 1940;61:179-83.
10. Ertel H. Der Verlauf der Vitamin-C-Prophylaxen im Frühjahr 1941. *Die Ernährung* 1941;6:269-73.
11. Cowan DW, Diehl HS, Baker AB. Vitamins for the prevention of colds. *JAMA* 1942;120:1268-71.
12. Ertel H. Über die Vitamin-C-Prophylaxen im Frühjahr 1942. *Die Ernährung* 1942;7:326-31.
13. Glazebrook AJ, Thomson S. The administration of vitamin C in a large institution and its effect on general health and resistance to infection. *Journal of Hygiene* 1942;42:1-19.
14. Oskam J. Onderzoek naar de waarde van vitamine C en chinine als prophylactica tegen "verkoudheidsziekten" [Dissertation]. Amsterdam, 1942.

15. Bergquist G. Om Profylaktiska medel mot s.k. förkylningssjukdomar. Svenska Lakartidningen 1943;40:2042-7.
16. Hopfengärtner. C-Vitaminprophylaxe bei Säuglingen. Monatsschr Kinderheilkd 1943;92:330-4.
17. Dahlberg G, Engel A, Rydin H. The value of ascorbic acid as a prophylactic against common colds. Acta Med Scand 1944;119:540-61.
18. Brown WMB, Mahoney F, Niedringhaus A, Locke A. Weather and susceptibility in relation to the spread of common cold; effect of ascorbic acid, in massive dosage, on duration. J Immunol 1945;50:161-79.
19. Cuendet O. Ernährungs- und Vitaminfragen im Saas-Tal. Praxis 1949;38:378-86.
20. Scheunert A. Der Tagesbedarf des Erwachsenen an Vitamin C. Internationale Zeitschrift für Vitaminforschung 1949;20:374-86.
21. Cowan DW, Diehl HS. Antihistaminic agents and ascorbic acid in the early treatment of the common cold. JAMA 1950;143:421-4.
22. Bartley W, Krebs HA, O'Brien JRP. Vitamin C requirement of human adults. Medical Research Council Special Report Series 1953;no.280.
23. Woolstone AS. Treatment of the common cold. Br Med J 1954;i:1290.
24. Bendel A. Redoxon als Prophylaktikum bei Erkältungskrankheiten. Ther Umsch Med Bibliograph 1955;12:87-8.
25. Franz WL, Sands GW, Heyl HL, Hanover NH. Blood ascorbic acid level in bioflavonoid and ascorbic acid therapy of common cold. JAMA 1956;162:1224-6.
26. Tebrock HE, Arminio JJ, Ossining NY, Johnston JH. Usefulness of bioflavonoids and ascorbic acid in treatments of common cold. JAMA 1956;162:1227-33.
27. Miegl H. Der akute Infekt der oberen Luftwege und seine Behandlung mit Vitamin C. Wien Med Wochenschr 1957;107:989-92.
28. Miegl H. Über die Anwendung von Vitamin C in der Hals-, Nasen-, Ohren-Heilkunde. Wien Med Wochenschr 1958;108:859-64.
29. Bessel-Lorck C. Erkältungsprophylaxe bei Jugendlichen im Skilager. Medizinische Welt 1959;(44):2126-7.
30. Barnes FE. Vitamin supplements and the incidence of colds in high school basketball players. N C Med J 1961;22:22-6.
31. Ritzel G. Kritische Beurteilung des Vitamins C als Prophylacticum und Therapeuticum der Erkältungskrankheiten. Helvetica Medica Acta 1961;28:63-8.
32. Shekhtman GA. On the significance of continuous addition of vitamin C to food in a military sector. Voennomed Zh 1961;3:46-9.
33. Dyllick H. Werkärztliche Erfahrungen mit Vitamin C. Medizinische Welt 1967;(17):1098-100.
34. Walker GH, Bynoe ML, Tyrell DAJ. Trial of ascorbic acid in the prevention of colds. Br Med J 1967;i:603-6.
35. Banks HS. Controlled trials in the early antibiotic treatment of colds. The Medical Officer 1968;70:7-10.
36. General Practitioner Research Group. Ineffectiveness of vitamin C in treating coryza. Practitioner 1968;200:442-5.
37. Regnier E. The administration of large doses of ascorbic acid in the prevention and treatment of the common cold. Rev Allergy 1968;22:948-56.
38. Wilson CWM, Loh HS. Ascorbic acid and upper respiratory inflammation. Acta Allergol 1969;24:367-70.
39. Anderson TW, Reid DBW, Beaton GH. Vitamin C and the common cold: a double-blind trial. Can Med Assoc J 1972;107:503-8.
40. Charleston SS, Clegg KM. Ascorbic acid and the common cold. Lancet 1972;i:1401-2.
41. Liljefors I. Vitamin C vid förkylningssjukdom. Lakartidningen 1972;69:3304-5.
42. Masek J, Neradilova M, Hejda S. Vitamin C and respiratory infections. Rev Czech Med 1972;18:228-35.
43. Schwartz AR, Togo Y, Hornick RB, Tominaga S, Gleckman RA. Evaluation of the efficacy of ascorbic acid in prophylaxis of induced rhinovirus 44 infection in man. J Infect Dis 1973;128:500-5.
44. Wilson CWM, Loh HS. Common cold and vitamin C. Lancet 1973;i:638-41.
45. Wilson CWM, Loh HS, Foster FG. The beneficial effect of vitamin C on the common cold. Eur J Clin Pharmacol 1973;6:26-32.
46. Wilson CWM, Loh HS, Foster FG. Common cold symptomatology and vitamin C. Eur J Clin Pharmacol 1973;6:196-202.

47. Anderson TW, Suranyi G, Beaton GH. The effect on winter illness of large doses of vitamin C. *Can Med Assoc J* 1974;111:31-6.
48. Carson M, Corbett M, Cox H, Pollitt N. Vitamin C and the common cold. *Br Med J* 1974;i:577.
49. Coulehan JL, Reisinger KS, Rogers KD, Bradley DW. Vitamin C prophylaxis in a boarding school. *N Engl J Med* 1974;290:6-10.
50. Sabiston BH, Radomski NW. Health problems and vitamin C in Canadian Northern military operations. Downsview, Ontario: Defence and Civil Institute of Environmental Medicine, 1974:report no.74-R-1012.
51. Anderson TW, Beaton GH, Corey PN, Spero L. Winter illness and vitamin C: the effect of a relatively low dosis. *Can Med Assoc J* 1975;112:823-6.
52. Clegg KM, MacDonald JM. l-Ascorbic acid and d-isoascorbic acid in a common cold survey. *Am J Clin Nutr* 1975;28:973-6.
53. Coulehan JL, Kapner L, Eberhard S. Vitamin C and upper respiratory illness in Navaho children: preliminary observations. *Ann N Y Acad Sci* 1975;258:513-22.
54. Karlowski TR, Chalmers TC, Frenkel LD, Kapikian AZ, Lewis TL, Lynch JM. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. *JAMA* 1975;231:1038-42.
55. Coulehan JL, Eberhard S, Kapner L, Taylor F, Rogers K, Garry P. Vitamin C and acute illness in Navajo schoolchildren. *N Engl J Med* 1976;295:973-7.
56. Elwood PC, Lee HP, Leger ASS, McLean Baird I, Howard AN. A randomized controlled trial of vitamin C in the prevention and amelioration of the common cold. *Br J Prev Soc Med* 1976;30:193-6.
57. Hejda S, Smola J, Masek J. Influence of physiological vitamin C allowances on the health status of miners. *Rev Czech Med* 1976;22:90-7.
58. Asfora J. Vitamin C in high doses in the treatment of the common cold. *Int J Vitam Nutr Res* 1977;(16 suppl):219-34.
59. Elwood PC, Hughes SJ, Leger ASS. A randomized controlled trial of the effect of vitamin C in the common cold. *Practitioner* 1977;218:133-7.
60. Gormly PJ. Megadose of ascorbic acid in an Antarctic expedition. *Br J Nutr* 1977;37:269-77.
61. Ludvigsson J, Hansson LO, Tibbling G. Vitamin C as a preventive medicine against colds in children. *Scand J Infect Dis* 1977;9:91-8.
62. Miller JZ, Nance WE, Norton JA, Wolen RL, Griffith RS, Rose RJ. Therapeutic effect of vitamin C. A co-twin control study. *JAMA* 1977;237:248-51.
63. Tyrrell DAI, Craig JW, Meade TW, White T. A trial of ascorbic acid in the treatment of the common cold. *Br J Prev Soc Med* 1977;31:189-91.
64. MacLean Baird I, Hughes RE, Wilson HK, Davies JEW, Howard AN. The effects of ascorbic acid on the occurrence of symptoms normally associated with the common cold. *Am J Clin Nutr* 1979;32:1686-90.
65. Pitt HA, Costrini AM. Vitamin C prophylaxis in marine recruits. *JAMA* 1979;241:908-11.
66. Carr AB, Einstein R, Lai LAC, Martin NG, Starmer GA. Vitamin C and the common cold. Using identical twins as controls. *Med J Aust* 1981;iii:411-2.
67. Carr AB, Einstein R, Lai LAC, Martin NG, Starmer GA. Vitamin C and the common cold: a second MZ Cotwin control study. *Acta Genet Med Gemellol (Roma)* 1981;30:249-55.
68. Martin NG, Carr AB, Oakeshott JG, Clark P. Co-twin control studies: vitamin C and the common cold. *Prog Clin Biol Res* 1982;103a:365-73.
69. Bancalari A, Seguel C, Neira F, Ruiz I, Calvo C. Valor profilactico de la vitamina C en infecciones respiratorias agudas del escolar. *Rev Med Chile* 1984;112:871-6.
70. Scheffler W. Über die betriebsärztliche Prophylaxe der Grippe- und Erkältungskrankheiten unter besonderer Berücksichtigung der Vitamin-C-Verabreichung [Dissertation]. Leipzig, 1966.
71. Bibile SW. Vitamin C in the prophylaxis and treatment of the common cold. *Proceedings of the Ceylon Association of Advanced Science* 1966;1(9):7.
72. Elliot B. Ascorbic acid: efficacy in the prevention of symptoms of respiratory infection on a polaris submarine. *International Research Communications System: Respiratory System* 1973;(5):12.
73. Kleijnen J, Riet ter G, Knipschild PG. Vitamine C en verkoudheid [Report]. Maastricht: Department of Epidemiology/Health Care Research, University of Limburg, 1988.

Special article

Vitamin E and cardiovascular disease

J. Kleijnen, P. Knipschild, and G. ter Riet

Department of Epidemiology/Health Care Research, University of Limburg, Maastricht, The Netherlands

Summary. Randomized clinical trials of the effects of Vitamin E on complaints of intermittent claudication and angina pectoris are reviewed, and their methodological shortcomings are considered. Mechanisms by which Vitamin E might act in cardiovascular disease are discussed. The evidence about the possible pharmacological actions of Vitamin E supports the need for randomized clinical trials. It is concluded that Vitamin E may have a positive effect, especially in intermittent claudication.

Key words: Vitamin E, angina pectoris; intermittent claudication

In the treatment of angina pectoris and intermittent claudication Vitamin E has had its proponents since the 1940s. Case histories, first from Vogelsang and the Shute brothers, led to more valid research by means of controlled clinical trials. The evidence regarding the alleged positive effects of Vitamin E in cardiovascular disease will be summarized here. Some of the proposed mechanisms by which Vitamin E acts in cardiovascular disease are discussed, and all the available randomized clinical trials are presented. Finally, the existing evidence is assessed and the need for further research is outlined.

Mechanisms of action

The most important function of Vitamin E in the body is to protect polyunsaturated fatty acids (PUFAs) from oxidation. PUFAs are highly susceptible to oxidation by endogenous free radicals, which are formed and take part in normal cell metabolism.

Absorbed exogenous free radicals, for instance from cigarette smoke, may cause oxidation, too. The cell damage that occurs during ischaemic periods in all tissues is probably caused by free radicals. In vitro experiments and in vivo animal studies have shown that Vitamin E protects mitochondria from the consequences of experimentally induced ischaemia [1]. Animal studies have shown that the most frequently occurring sign in Vitamin E deficiency is myopathy, including myopathy of the heart muscle [2]. In patients with Vitamin E deficiency, abnormal erythrocytes and sometimes anaemia are also found and in them Vitamin E increases the life-span of red blood cells [3, 4]. Deformability of red blood cells may be enhanced by Vitamin E, since PUFAs incorporated in the membranes are protected from oxidation. Finally, Vitamin E inhibits platelet aggregation [5]. Other, more controversial hypotheses about the actions of Vitamin E are its ability to lower the blood cholesterol level [6-9] and to stimulate the formation of collateral vessels [10]. An improvement in physical working capacity on Vitamin E supplementation is also subject to controversy [11, 12].

How can this information be translated into mechanisms of action? Vitamin E might improve tolerance to ischaemia that occurs during exercise, if it were to eliminate free radicals. It might also influence the process of atherosclerosis by stopping further deterioration, but this will be difficult to prove. It was recently shown that patients with ischaemic heart disease and patients with peripheral arterial disease had higher plasma lipid peroxide concentrations than controls [13]. Inhibition of peroxidation by Vitamin E might favourably influence the balance between peroxidative damage and the repair mechanisms of the body. Finally, an influence

on platelet aggregation and its effects on red blood cells, both improving the blood flow, could account for some beneficial effects.

Methods

Articles on experiments were found by a MEDLINE computer search (1963-1988) for "Vitamin E and clinical trials", and by extensively checking the references in the publications we already possessed.

This section contains comments on methodological criteria for assessment of the controlled trials.

Prognostic comparability is not guaranteed by randomization, especially when small numbers of patients are used. Prestratification of important modifiers of the treatment effect, such as severity of disease, is important, and possible confounding factors should be carefully measured. To ensure the same intervention for all patients, normal treatments, like exercise, weight loss and stopping smoking, should be continued in both treatment groups and Vitamin E or an indistinguishable placebo should be given.

In practice, for trials using patients with intermittent claudication, the subjective perception of pain, disability and the walking distance by the patient should be considered as the relevant outcome phenomena. 'Objective' measurements, like the ankle-arm blood pressure index, the walking distance until onset of pain on a standard tread mill, and the occurrence of vascular surgery, amputation or death could be measured. The follow-up period should be at least three months, although some authors have recommended one or more years. Both investigators and patients should be blinded and compliance must be assessed.

In trials using patients with angina pectoris, valid measurements of the treatment effect should be ascertained in a similar way. In our opinion trials with fewer than about 50 patients in each group have limited value, since small but clinically relevant beneficial effects might remain undetected, and prognostic comparability might not be achieved, even in spite of an optimal study design.

Results

For interpretation of the doses, it is necessary to consider the Vitamin E activity of the various tocopherols. The Vitamin E activity of 1 mg synthetic d- α -tocopherol acetate is equivalent to 1 IU of Vitamin E. D1- α -tocopherol has a potency of

1.1 IU/mg. The activity of naturally occurring d- α -tocopherol is 1.49 IU/mg, and of its acetate 1.36 IU/mg [14].

Angina pectoris

Six controlled trials of Vitamin E in angina pectoris were found [15-22], of which only 2 were randomized and double-blind [17-20]. Both used only some 20 patients in each treatment group.

Rinzler [17, 18] gave 200 mg synthetic alpha-tocopherol acetate for 2 weeks and thereafter 300 mg daily or placebo to ambulant patients with chronic chest pain and with arteriosclerotic and/or hypertensive heart disease. Alpha-tocopherol was administered for 10 to 20 weeks, with an average of 16 weeks. The number of patients was further reduced, because 25% of the patients had "somatic chest pain" with no effort angina.

Anderson [19, 20] gave no less than 3200 IU Vitamin E daily or an indistinguishable placebo for 9 weeks to patients whose angina was reasonably stable. During the study, 10 patients had tried Vitamin E at some time on their own initiative, and the consumption of other medications, such as nitroglycerin, was not equally divided at base line between the index and the referent groups.

No or only small differences in favour of the Vitamin E-treated groups were found in the pain score and other outcome measurements in both trials. The most important limitation of both trials is the small number of patients used. Of the other 4 trials, especially Gillilan's [21] deserves attention. In a double-blind cross-over study 1600 IU d- α -tocopherol succinate or placebo was given daily. The trial periods of 6 months were separated by a 2 months wash-out period. In 52 patients with angina pectoris no differences were found for exercise capacity, left ventricular function or frequency of chest pain.

Intermittent claudication

There were 5 controlled trials of Vitamin E in intermittent claudication (Table 1; 23-28). Except for Boyd's study, which was probably not randomized and single-blind, all experiments were randomized and double-blind. The trials of Ratcliffe [29] and Haeger [10, 30, 31] have been excluded since they were nonexperimental cohort studies. Both showed a favourable outcome in groups treated with Vitamin E.

Table 1. Trials assessing the effect of Vitamin E on intermittent claudication

Author (year)	Number of patients/controls	Daily dosage/follow-up period	Result for pain, Vit. E/placebo	Result for walking distance, Vit. E/placebo
Hamilton (1953)	20/21	450 IU wheat germ oil/12 weeks	-9.7%/+8.3% of pretreatment average	21/19 circuits (pretreatment 19/18)
Livingstone (1958)	20/20	600 mg (?) /40 weeks	not given	39/31.5 circuits (pretreatment 24/29)
Williams (1962)	16/17	1600 mg d- α -tocopherol acetate/at least 3 months	not measured	478/265 yards (pretreatment 223/266)
Boyd (1963)	17/16	400 mg dl- α -tocopherol/13 weeks	distance to onset pain 330/240 yards (pretreatment 133/184)	distance to halting 661/569 yards (pretreatment 532/719)
Westheim (1975)	40/40	900 mg for 2 months, then 300 mg for 6 months d- α -tocopherol/8 months	not measured	number of times standing on tiptoe 58/48 (pretreatment 41/36)

The study of Hamilton et al. was criticized for its use of wheat germ oil as the source of Vitamin E [23]. In the early 1950s the content of tocopherol in germ oil was often considerably less than the declared value, due to inaccurate measurement. Marks stated that the 450 mg used by Hamilton might represent a true α -tocopherol content of about 40 mg [32].

The other controlled trials show positive results in the Vitamin E-treated groups. Virtually no side effects were reported, even in the trials in which extremely high doses of Vitamin E were used. Again, the major limitation is the small study size. No indication of the magnitude of the beneficial effect can be given.

Pooling of the data appears to be useless, since there were great differences in the doses used, in the disease spectra of the patients in the experiments and in the outcome measurements.

Discussion

All the experiments on intermittent claudication have relied on patients who were not eligible for surgery. We believe that patients who are eligible for surgery should not be treated with Vitamin E alone, unless its efficacy is proven in pragmatic trials comparing both treatments. Precise descriptions of prognostic factors are essential to detect subgroups who might react positively to Vitamin E supplements. In

the trials discussed above, this information was incomplete. Most patients were described according to Boyd's classification of complaints [33]. Exclusion of patients with other causes of pain, e.g. osteoarthritis, was only mentioned once. The doses of Vitamin E ranged from 300 to 3600 IU daily. In no experiment were several dosages employed to assess the optimal level, although it could be argued that for efficiency this should only be done after a positive effect has been demonstrated. In only two experiments were post-treatment effects assessed. The most important weakness, however, as stated before, was the small number of patients enrolled.

What is the effect of shortcomings of trials on interpretation of the results? Unless it can be argued that some systematic bias to positive results would occur, and this does not appear to be the case, it would be expected that positive and negative trials would occur with the same frequency. A surplus of positive trials is evidence then of a positive effect of Vitamin E. The experiments on Vitamin E in the treatment of intermittent claudication tend towards positive results, while the results of Vitamin E supplementation in angina pectoris are less convincing. The latter could well be the result of a lack of power, but a large effect seems unlikely. It is remarkable that the most recent trial found was published in 1977, twelve years ago! Why haven't these trials been confirmed by experiments with larger numbers of patients? Or were such trials not published or stopped at an early phase because of (preliminary) negative results? Publication bias is an important complicating factor (i.e. negative trials being rejected or not submitted for publication) and it can influence the results of a review. Estimation of the extent of publication bias is difficult, and it is also possible that some publications were missed.

Conclusion

The evidence of the effects of Vitamin E on cardiovascular disease, with the limitations stated above in mind, is slightly in favour of a positive effect of Vitamin E, especially in the treatment of intermittent claudication. The conclusion must be that larger well-designed double-blind trials are necessary to confirm or reject the results of the existing studies, to assess optimal dosage levels and to detect subgroups of patients with peripheral arterial occlusive disease in whom Vitamin E treatment might be beneficial.

Acknowledgement. The review was supported by a grant from the Dutch Ministry of Welfare, Public Health and Cultural Affairs.

References

1. Ferrari R, Visioli O, Guarnieri C, Caldarera M (1983) Vitamin E and the heart: Possible role as antioxidant. *Acta Vitaminol Enzymol* 5: 11-22
2. Mächlin LJ (1984) Vitamin E. In: Mächlin LJ (ed) *Handbook of vitamins*. Marcel Dekker, New York, pp 99-145
3. Leonard PJ, Losowsky MS (1971) Effect of alpha-tocopherol administration on red cell survival in vitamin E-deficient human subjects. *Am J Clin Nutr* 24: 388-393
4. Farrell PM, Bieri JG, Fratantoni JF, Wood RE, di Sant' Agnese PA (1977) The occurrence and effects of human vitamin E deficiency. *J Clin Invest* 60: 233-241
5. Steiner M, Mower R (1982) Mechanism of action of vitamin E on platelet function. *Ann NY Acad Sci* 393: 289-299
6. Hermann WJ, Ward K, Faucett J (1979) The effect of tocopherol on high-density lipoprotein cholesterol. *Am J Clin Pathol* 72: 848-852
7. Howard DR, Rundell CA, Batsakis JG (1982) Vitamin E does not modify HDL-cholesterol. *Am J Clin Pathol* 77: 86-89
8. Stampfer MJ, Willett W, Castelli WP, Taylor JO, Fine J, Hennekens CH (1983) Effect of vitamin E on lipids. *Am J Clin Pathol* 79: 714-716
9. Cloarec MJ, Perdriset GM, Lamberdiere FA, Colas-Belcour JF, Sanzies JP, Neufeld HN, Goldbourt U (1987) α -Tocopherol: effect on plasma lipoproteins in hypercholesterolemic patients. *Isr J Med Sci* 23: 869-872
10. Haeger K (1982) Long-term study of α -tocopherol in intermittent claudication. *Ann NY Acad Sci* 393: 369-375
11. Shephard RJ (1983) Vitamin E and athletic performance. *J Sports Med* 23: 461-470
12. Simon-Schnass I, Pabst H (1988) Influence of vitamin E on physical performance. *Int J Vit Nutr Res* 58: 49-54
13. Stringer MD, Görög PG, Freeman A, Kakkar VV (1989) Lipid peroxides and atherosclerosis. *Br Med J* 298: 281-284
14. Recommended Dietary Allowances (1980) National Academy of Sciences, Washington, DC, p 64
15. Makinson DH, Oleesky S, Stone RV (1948) Vitamin E in angina pectoris. *Lancet* 1: 102
16. Donegan CK, Messer AL, Orgain ES, Ruffin JM (1949) Negative results of tocopherol therapy in cardiovascular disease. *Am J Med Sci* 217: 294-299
17. Travell J, Rinzler SH, Bakst H, Benjamin ZH, Bobb AL (1949) Comparison of effects of alpha-tocopherol and a matching placebo on chest pain in patients with heart disease. *Ann NY Acad Sci* 52: 345-353
18. Rinzler SH, Bakst H, Benjamin ZH, Bobb AL, Travell J (1950) Failure of alpha tocopherol to influence chest pain in patients with heart disease. *Circulation* 1: 288-293
19. Anderson TW (1974) Vitamin E in angina pectoris. *Can Med Assoc J* 110: 401-406
20. Anderson TW, Reid W (1974) A double-blind trial of vitamin E in angina pectoris. *Am J Clin Nutr* 27: 1174-1178
21. Gillilan RE, Mondell B, Warbasse JR (1977) Quantitative evaluation of vitamin E in the treatment of angina pectoris. *Am Heart J* 93: 444-449
22. Toone WM (1973) Effects of vitamin E: Good and bad. *N Engl J Med* 289: 979-980
23. Hamilton M, Wilson GM, Armitage P, Boyd JT (1953) The treatment of intermittent claudication with vitamin E. *Lancet* 1: 367-370
24. Livingstone PD, Jones C (1958) Treatment of intermittent claudication with vitamin E. *Lancet* 2: 602-604
25. Williams HTG, Clein LJ, Macbeth RA (1962) Alpha-tocopherol in the treatment of intermittent claudication. A preliminary report. *Can Med Assoc J* 87: 538-541
26. Clein LJ, Williams HTG, Macbeth RA (1962) The use of alpha-tocopherol in the treatment of intermittent claudication. *Can Med Assoc J* 86: 215-216
27. Boyd AM, Marks J (1963) Treatment of intermittent claudication. A reappraisal of the value of α -tocopherol. *Angiology* 14: 198-208
28. Westheim AS, Brox D, Selvaag AW (1975) D- α -tokoferol ved claudicatio intermittens. En klinisk undersøkelse. *Tidsskr Nor Laegeforen* 95: 13-15
29. Ratcliffe AH (1949) Vitamin E in intermittent claudication. *Lancet* 2: 1128-1131
30. Haeger K (1968) The treatment of peripheral occlusive arterial disease with a α -tocopherol as compared with vasodilator agents and antiproteolytic (dicumarol). *Vasc Dis* 5: 199-213
31. Haeger K (1974) Long-time treatment of intermittent claudication with vitamin E. *Am J Clin Nutr* 27: 1179-81
32. Marks J (1962) Critical appraisal of the therapeutic value of α -tocopherol. *Vitam Horm* 20: 573-598
33. Boyd AM (1951) *Proc Roy Soc Med* 44: 985-989

Received: March 3, 1989
accepted in revised form: June 21, 1989

Dr. J. Kleijnen
Department of Epidemiology/Health Care Research
University of Limburg
Postbox 616
NL-6200 MD Maastricht
The Netherlands

REVIEW

Niacin and Vitamin B6 in Mental Functioning: A Review of Controlled Trials in Humans

Jos Kleijnen and Paul Knipschild

Fifty-three controlled trials of the effects of niacin, vitamin B6, and multivitamins on mental functions are reviewed. The results are interpreted with emphasis on the methodological quality of the trials. It turns out that virtually all trials show serious shortcomings: in the number of participants, the presentation of baseline characteristics and outcomes, and the description of changes in concomitant treatments. Only in autistic children are some positive results are found with very high dosages of vitamin B6 combined with magnesium, but further evidence is needed before more definitive conclusions can be drawn. For many other indications (hyperactive children, children with Down's syndrome, IQ changes in healthy schoolchildren, schizophrenia, psychological functions in healthy adults and geriatric patients) there is no adequate support from controlled trials in favor of vitamin supplementation.

Introduction

In the last four decades, megadoses of vitamins have been tested for many conditions. Trials on the effects of vitamins in schizophrenia were followed by similar trials on mental diseases in children: autism, Down's syndrome, hyperactivity, and mental retardation. Trials have also been conducted on healthy adults and elderly people on memorization, psychological functioning, and dementia. Recently, four trials were published on the effects of megadoses of vitamins on intelligence in healthy children. More than a decade ago, after extensive discussions in the literature, most physicians were skeptical about the beneficial effects of vitamins in psychiatry. Many controlled trials have been performed since then. We found no less than 30 publications of controlled trials since 1980. Common denominator in all these trials was the use of niacin or vitamin B6 alone, or in combination with other vitamins and minerals. We confined this review to niacin and vitamin B6 because these are the most frequently used vitamins in controlled trials and there is a biological rationale for their use. This review article briefly describes the biological rationale for the use of niacin and vitamin B6 in mental ailments, and summarizes the results of controlled trials in humans. The methodological quality of the trials will be emphasized.

From the Department of Epidemiology/Health Care Research, University of Limburg, The Netherlands.
Address reprint requests to J. Kleijnen, M.D., Department of Epidemiology/Health Care Research, University of Limburg,
P.O. Box 616, 6200 MD Maastricht, The Netherlands.
Received September 10, 1990; revised November 2, 1990.

Mechanisms

There are two forms of niacin: nicotinic acid and nicotinamide. In the Index Medicus terminology, nicotinic acid is called niacin, but often niacin is used as a collective name for all forms of vitamin B3 (Anonymous 1987). Niacin acts as a hydrogen carrier in its coenzyme forms nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP) (e.g., in the Krebs cycle). Hundreds of enzymes (dehydrogenases), involved in the synthesis and/or degradation of carbohydrates, fatty acids, and amino acids, are dependent on NAD or NADP as a coenzyme (Friedrich 1988). Nicotinic acid is known as a peripheral vasodilator and there have been claims that it increases cerebral circulation. Also, there have been controversies over a possible influence of niacin on an abnormal methylation of neurotransmitters in schizophrenic subjects (Nestoros et al 1977).

There are six biologically active forms of vitamin B6: pyridoxine, pyridoxal, pyridoxamine, and their respective phosphates. In the form of pyridoxal phosphate, vitamin B6 serves as a coenzyme for a wide variety of enzymes of amino acid metabolism. Human beings cannot synthesize vitamin B6 *de novo*, contrary to the formation of nicotinic acid from tryptophan, in which vitamin B6 also serves as a coenzyme. However, the six biologically active forms of vitamin B6 can be converted into each other. Vitamin B6 is involved in crucial steps of the synthesis of several neurotransmitters, for instance, in the decarboxylation of glutamic acid into gamma amino butyric acid (GABA). In humans, vitamin B6 is transported into the central nervous system by the cerebrospinal fluid (Friedrich 1988; Hodges 1980; Lipton et al 1987). In patients with homocystinuria resulting from cystathionine β -synthase deficiency, mental retardation and behavioral disturbances are often found. It has been shown that megadoses of vitamin B6 (25-500 mg/day) produce a fall in plasma and urinary methionine and homocystine and an increase in cytosine concentration in body fluids in approximately half of patients who take them (Braunwald et al 1987).

Aberrations of the metabolism of niacin, vitamin B6, and tryptophan have been described in several mental diseases, and in both vitamin B6 deficiencies (convulsions and degeneration of peripheral nerves) and niacin deficiencies (dementia, hallucinations, psychosis) neurological symptoms can be found. Some biological plausibility of (positive) effects on mental functioning can easily be distilled from the role of these vitamins in neurotransmitter metabolism and from the mechanisms mentioned above. Although there are many unsolved parts in this puzzle, for example, the influence of previous or concomitant medication, there seems to be some rationale for controlled clinical trials.

Materials and Methods

Identification and Selection of Relevant Studies

Human trials were found by checking references extensively, by personal communication with authors of experiments, and by means of a MEDLINE CD-ROM search from 1983 to 1990 (key words: pyridoxine-therapeutic-use, niacin-therapeutic-use, nicotinamide-therapeutic-use, vitamins-therapeutic-use). Further trials were found by searching the Psychiatry section of Excerpta Medica from 1965 onward, using the following key words: nicotinamide, nicotinic acid, pyridoxine, and vitamin(s) (B3) (B6). Checking the Index Medicus from 1965 onward yielded no additional articles. Trials were only considered if a parallel control group was included, if niacin or vitamin B6 alone or in combination

with other vitamins/minerals was used, and if correlates of mental state were assessed. We excluded trials in which niacin or vitamin B6 (or analogues) combined with pharmaceutical agents were tested. Consequently, trials such as Bumbalo et al's (1964) in children with Down's syndrome, assessing the effect of 48 medications including vitamins and minerals but also pharmaceutical agents were excluded. Another example of exclusion are trials in which xanthinol nicotinate was tested (a combination of nicotinic acid and the theophylline derivate xanthinol).

Assessment of Methodological Quality of Published Studies

To explore the possibility that an increasing likelihood of bias (an increasing number of methodological shortcomings) is reflected in the results of the trials, the following criteria for a methodological assessment of the trials were established in advance: (1) well-described diagnosis (DSM diagnosis or diagnostic criteria; description of symptoms, duration, severity, and therapy); (2) at least 50 patients per group; (3) prestratification (matching) on relevant prognostic factors; (4) randomization; (5) presentation of relevant baseline characteristics; (6) patients blinded; (7) evaluator blinded; (8) effect measurement well described (reproducible); (9) assessment of other treatments or vitamins; (10) presentation of the results in such a manner that the analysis can be checked by the reader (depending on the outcome measurement: at least the mean and standard deviation, standard error, or confidence interval per group have to be mentioned, or, if rates or proportions are used, the number of patients with a certain outcome).

A complicating factor in blinding the patients and the evaluator is flushing, which is often seen after the intake of large doses of nicotinic acid. In some trials this problem was partially solved by giving all participants nicotinic acid initially, followed by a gradual replacement by placebo. Also, slow release tablets have been used. Another solution might be the exclusion of patients who develop severe flushing initially, before randomization takes place.

Assessment of articles using these criteria offers an indication of the methodological quality of the trials. This quality will be an important factor in weighing the conclusions of different trials and, of course, it also bears on the impact on the reader's opinion of all the evidence presented. The function of the total score (Tables 1 and 2) is only a quick reference to the methodological quality, and does not mean that we presuppose some sort of numerical equality among the different criteria. It could be argued that other criteria should be used for the methodological assessment and that this kind of assessment is rather subjective. However, we have selected well-established methodological criteria, and our assessment can be checked by the reader (Tables 1 and 2). Also, the reader could use other criteria to see whether substantial changes would occur in the methodological assessment. We do not think that this will be the case.

Results

We found reports of 53 trials, which include 5 trials in autistic children, 7 trials in hyperactive or learning disabled children, 10 trials in children with Down's syndrome, 4 trials on IQ changes in healthy schoolchildren, 2 trials in learning disabled adults, 18 trials in schizophrenic patients, 5 trials concerned with the effects on memorization and psychological functioning in healthy people, and finally 2 trials in geriatric patients.

Methodological assessment showed that frequently shortcomings existed in the numbers

Table 1. The Methodological Score Assigned to the Experiments in Children and Young Adults

First author, year	Result	A	B	C	D	E	F	G	H	I	J	TOT
Autism												
Rimland '78	Sign.	-	-	-	+	±	+	+	±	-	±	4.5
Barthelemy '80	=	+	-	-	+	-	+	+	+	-	-	5
Barthelemy '80	Sign.	+	-	-	+	-	+	+	+	-	-	5
Martineau '85	0	+	±	-	+	-	+	+	+	-	-	5.5
Jonas '84	Sign.	+	-	-	-	-	+	+	+	-	+	5
Hyperactivity & learning disability												
Kerschner '77	0	±	-	+	+	+	+	+	+	±	+	8
Arnold '78	0	+	-	-	+	±	+	+	+	-	-	5.5
Coleman '79	Pos.	+	-	-	+	-	+	+	+	-	-	5
Maseck '80	Sign.	-	-	-	-	-	-	-	±	-	+	1.5
Colgan '84	Sign.	-	-	+	-	+	+	+	+	-	+	6
Haslam '84	Neg.	+	-	-	+	-	+	+	+	-	-	5
Boggs '85	Pos.	+	-	-	+	+	+	+	+	-	-	6
Down's syndrome												
Pueschel '80	0	+	-	-	+	+	+	+	+	+	+	8
Harrell '81	Sign.	-	-	+	-	+	+	+	+	-	+	6
Bennett '83	0	+	-	+	+	+	+	+	+	+	+	9
Coburn '83	0	-	-	-	-	+	+	+	±	-	+	4.5
Weathers '83	0	+	-	±	+	+	+	+	±	-	+	7
Ellman '84	0	±	-	+	-	+	+	+	+	-	+	6.5
Smith '84	0	+	±	+	+	+	+	+	+	±	+	9
Coleman '85	0	+	-	-	+	-	+	+	+	-	-	5
Bidder '89	Neg.	+	-	+	+	+	+	+	+	-	+	8
Menolascino '89	0	-	-	-	+	-	+	+	+	-	-	4
Schoolchildren (IQ changes)												
Benton '88	Sign.	+	±	+	+	-	+	+	+	-	±	7
Naismith '88	0	±	+	+	-	+	+	+	+	-	+	7.5
Crombie '90	0	+	±	+	+	+	+	+	+	-	±	8
Benton '90	Pos.	±	+	±	-	-	+	+	-	-	-	4
Schizophrenia												
Greenbaum '70	0	±	-	-	+	-	+	+	+	-	-	4.5

+: item is scored; -: item is not scored; ±: item is partly scored (25-50 patients in category B); TOT: total number of items scored; Sign.: index group fared better (author stated that result was significant); Pos.: positive trend (or significant effect for only some of the effect measurements); 0: no difference; =: no difference compared with control group treated with magnesium; Neg.: control group fared better; A: well-described diagnosis (DSM diagnosis or diagnostic criteria; description of symptoms, duration, severity, and therapy); B: at least 50 patients per group; C: prestratification (matching) on relevant prognostic factors; D: randomization; E: presentation of relevant baseline characteristics; F: patients blinded; G: evaluator blinded; H: effect measurement well described (reproducible); I: use of other treatments or vitamins has been assessed; J: presentation of the results in such a manner that the analysis can be checked by the reader. Depending on the outcome measurement: at least the mean(s) and standard deviation, standard error, or confidence interval per group have to be mentioned, or the number of patients with a certain outcome (e.g., if rates or proportions are used).

of participants, the presentation of baseline characteristics of the respective groups, the description of changes in other therapies (medication, vitamins, behavior therapy), and the data presentation. Only 12 experiments showed flaws on less than four of the items, indicating that the overall quality of the experiments was low. Interestingly, eight studies with the highest scores showed no positive effects of supplementation. Note that we could only score on items that were actually mentioned in the publication, that is, randomization was not scored if the trial was double-blind, but only if randomization was specifically mentioned.

Table 2. The Methodological Score Assigned to the Experiments in Adults

First author, year	Result	A	B	C	D	E	F	G	H	I	J	TOT
Learning disability												
Ellis '83	0	-	-	+	-	+	+	+	+	-	+	6
Chanowitz '85	0	-	-	+	-	+	+	+	+	-	+	6
Schizophrenia												
O'Reilly '55	Pos.	±	-	-	-	±	-	-	+	-	-	2
Hoffer '57	Sign.	+	-	-	+	-	+	+	-	±	±	5
Ashby '60	0	-	±	+	-	-	+	+	+	-	-	4.5
Denson '62	Sign.	±	-	-	-	+	+	+	±	+	+	6
Hoffer '62	Sign.	+	±	-	+	+	+	+	-	±	-	6
Ananth '70	Sign.	±	-	-	-	-	+	+	+	-	-	3.5
Ramsay '70	0	±	-	-	+	+	+	+	+	+	±	7
Valley '71	0	±	-	+	-	-	+	+	+	-	-	4.5
Ananth '72	0	±	-	-	+	-	+	+	+	+	-	5.5
McGrath '72	0	+	+	-	+	-	+	+	±	-	±	6
Ananth '73	Pos.	±	-	-	+	-	+	+	+	+	-	5.5
Altman '73	0	-	+	-	+	-	+	+	±	-	-	4.5
Wittenborn '73	0	±	±	-	+	+	+	+	+	+	±	7.5
Sehdev '74	Pos.	±	-	-	-	±	+	-	+	-	-	3
Ban '77	Neg.	-	-	-	-	-	+	+	+	-	-	3
Deutsch '77	Neg.	-	-	-	+	-	+	+	+	+	-	5
Dries '78	Sign.	-	±	-	-	+	+	+	+	-	-	4.5
Memorization & psychological functioning (healthy adults)												
Molimard '80	Neg.	±	±	-	+	-	+	+	+	-	±	5.5
Molimard '80	Neg.	±	-	+	+	-	+	+	+	-	±	6
Chomé '84	0	-	±	-	+	-	+	+	+	-	-	4.5
Loriaux '85	Sign.	±	±	+	-	-	+	+	+	-	-	5
Heseker '90	0	+	+	-	+	-	+	+	±	-	-	5.5
Geriatric patients												
Lu '71	0	-	-	-	+	±	+	+	+	-	-	4.5
Blass '88	0	+	-	-	+	-	+	+	+	-	-	5

Abbreviations same as Table 1.

The recommended dietary allowances (RDA) for niacin and vitamin B6 are about 15 and 2 mg, respectively. Although in some trials moderate dosages were used, Table 3 shows that mostly very high dosages were administered, up to 600 and 1500 times the RDA, respectively. In some studies, a fairly heterogeneous study population was used, including patients with several diagnoses, that is, a group of institutionalized patients. These studies can be identified in Tables 3 and 4. The results in these tables reflect the conclusions as given by the original author(s) of the trials.

In autism, consistently positive results have been reported of supplementation with pyridoxine combined with magnesium (Barthelemy et al 1980; Jonas et al 1984). Pyridoxine alone showed contradictory results (Rimland et al 1978; Martineau et al 1985). The doses of vitamin B6 in these studies was approximately three times higher than in the other indications. Although all trials are rather small, the results are encouraging and further larger trials are warranted.

In seven trials, the effects of vitamins in children with hyperactivity or learning disability were assessed. These diagnoses seem to overlap to some extent in the groups of participants of the trials. Table 3 shows that the number of patients included was very

Table 3. Characteristics of the Intervention and the Results of the Experiments (Children and Young Adults)

First author	Patients/ controls	NA	NM	B6	Other vitamins/ minerals	Duration (mo)	Result
Autism							
Rimland '78	16 cross-over			3000		1	Sign.
Barthelemy '80	16 cross-over			30/kg	Mg	<1	=
Barthelemy '80	21 cross-over			30/kg	Mg	<1	Sign.
Martineau '85	37 cross-over			30/kg		<1	0
Jonas '84	8 cross-over			1000	Mg	1.5	Sign.
Hyperactivity & learning disability							
Kerschner '77	10/10		3000	750	B5, C	6	0
Arnold '78	15/16	1000		100	Multi	<1	0
Coleman '79	6 cross-over			30/kg	B5, C	1	Pos.
Maseck '80	11/11		??	??	Multi	5	Sign.
Colgan '84	16/16		1000	450	Multi	5	Sign.
Haslam '84	7 cross-over		3000	600		1	Neg.
Boggs '85	6 cross-over	400		100	Multi	1	Pos.
Down's syndrome							
Pueschel '80	23/20			10/kg		36	0
Harrell '81	5/11 #		750	350	Multi	4	Sign.
Bennett '83	10/10		750	350	Multi	8	0
Coburn '83	20/18 #		750	350	Multi	5	0
Weathers '83	24/23		750	350	Multi	4	0
Ellman '84	10/10 #		750	350	Multi	6	0
Smith '84	28/28		750	350	Multi	8	0
Coleman '85	10/9			35/kg		36	0
Bidder '89	19 cross-over	570		350	Multi	3	Neg.
Menolascino '89	11/13 #		750	350	Multi	8	0
Healthy schoolchildren (IQ changes)							
Benton '88	30/60	50		10	Multi	8	Sign.
Naismith '88	77/77	50		10	Multi	1	0
Crombie '90	42/44	50		10	Multi	7	0
Benton '90	84/83		20	2	Multi	5	Pos.
Schizophrenia							
Greenbaum '70	17/24		44/kg			6	0

NA: nicotinic acid (highest dosage in mg); NM: nicotinamide (highest dosage in mg); B6: pyridoxine (highest dosage in mg); #: also other diagnoses in study population; ?: exact number of patients and controls not mentioned; ??: dosage not mentioned.

small, and in one publication even the dosage of the vitamins was not mentioned (Maseck 1980). The results are contradictory. Two trials in retarded adults showed no positive effects (Table 4) (Ellis and Tomporowski 1983; Chanowitz et al 1985).

In children with Down's syndrome no positive effects of pyridoxine or multivitamin preparations were found in 9 of 10 studies. Although four studies also included children with other diagnoses, the weight of the evidence is against supplementation with (multi) vitamins in Down's syndrome, and some large, well-performed studies with positive results would be needed to change this status.

Two trials in which a multivitamin preparation was tested on healthy children showed positive IQ changes in the index group (Benton and Roberts 1988; Benton and Buts

Table 4. Characteristics of the Intervention and the Results of the Experiments (Adults)

First author	Patients/controls	NA	NM	B6	Other vitamins/ minerals	Duration (mo)	Result
Learning disability							
Ellis '83	19/21		750	350	Multi	7	0
Chanowitz '85	37?		750	350	Multi	4	0
Schizophrenia							
O'Reilly '55	11/43	3000				2	Pos.
Hoffer '57, '62	10;11/9	3000	3000			1	Sign.
Ashby '60	39 cross-over	6300	6300			2	0
Denson '62	17/19		3000			1	Sign.
Hoffer '62	39/43	3000				1	Sign.
Ananth '70	20 cross-over	3000				<1	Sign.
Ramsay '70	10;10/10	3000	3000			6	0
Vallely '71	18;18/18	3000	3000			6	0
Ananth '72	9;10/11	9000	9000			24	0
McGrath '72	89/95		3000			12	0
Ananth '73	30?	3000		75		11	Pos.
Altman '73	69/63 #	50	5		Multi	1	0
Wittenborn '73	47/28	3000				24	0
Sehdev '74	14/13	3000				3	Pos.
Ban '77	20?	9000		150	Multi	5.5	Neg.
Deutsch '77	10;10/10	3150	3150			11	Neg.
Dries '78	36/37 #		30	10	Multi	3	Sign.
Memorization & psychological functioning (healthy adults)							
Molimard '80	58? students			500		<1	Neg.
Molimard '80	30? obese adults			1000		<1	Neg.
Chomé '84	34 cross-over	10		2	Multi	1	0
Loriaux '85	29/34	142				2	Sign.
Heseker '90	498/583	20		5	Multi	2	0
Geriatric patients							
Lu '71	20/40 various diagnoses	45				3	0
Blass '88	11 cross-over, Alzheimer's disease		750			3	0

Abbreviations same as in Tables 1, 2, and 3.

1990), whereas two other trials in which the same preparation was tested showed no difference between the index and control groups (Naismith et al 1988; Crombie et al 1990).

Only five of 18 trials showed significant positive results in schizophrenic patients. Table 2 shows that both trials with positive and with negative outcomes had several methodological shortcomings, and Table 4 shows that the weight of the evidence is not in favor of vitamin supplementation. In two trials, the placebo group even fared better (Ban et al 1977; Deutsch et al 1977).

In a small trial, no positive effects were found from nicotinamide in patients with Alzheimer's disease (Blass et al 1988).

In healthy adults, one study showed positive effects from nicotinic acid (Loriaux et al 1985), whereas two studies showed no positive effects from pyridoxine in tests assessing memorization (Molimard et al 1980). Neither were positive effects found in two studies

assessing the effects of a multivitamin preparation on psychological parameters in healthy adults who were found to have low vitamin levels (Chomé et al 1984; Heseker et al 1990).

The scales used for the assessment of the effects have generally been described adequately. In autism, tests assessing several autistic symptoms were used to measure the treatment effects. The Wechsler Intelligence Scale for Children (WISC) and several behavior rating scales were used to measure treatment effects in children with Down's syndrome, hyperactivity, and learning disabilities. The Brief Psychiatric Rating Scale and many other corresponding scales were used for the effect measurement in schizophrenic patients.

The most frequently reported adverse effects were flushing when nicotinic acid was used, and gastrointestinal complaints when nicotinamide or pyridoxine were used.

Discussion

The knowledge of the function of vitamins in the human body has expanded rapidly in the last decades. Although many questions remain, there seems to be a rationale for the proposition that high dosages of niacin and vitamin B6 might have beneficial effects. The trials in which moderate amounts were used may have a different rationale for therapy: the elimination of subclinical deficiencies might have beneficial effects. Whenever claims of possible therapeutic effects of a new treatment are supported by both a plausible explanation of the mechanism and case reports with positive results, randomized clinical trials are nowadays used as the criteria for establishing whether this new treatment is of any value. If carefully executed, these trials give the best possible estimation of the effects of an intervention.

Tables 1 and 2 show that virtually all published controlled trials have several methodological flaws. The most frequently encountered were the small number of participants, the lack of ascertainment of prognostic comparability of the index and control groups, inadequate registration of other (self-administered) therapies, and incomplete presentation of the results. The assessment of the quality of the trials is arbitrary. However, the presentation of the assessment in the tables, combined with descriptions of the literature search methods and the inclusion criteria, enables the reader to check our conclusions.

The overall number of trials with negative and positive outcomes (33 and 20, respectively) indicates that bias as a result of methodological shortcomings probably has not induced an overoptimistic view of the state of affairs; whether the opposite is true is hard to say. It is also reasonable to assume that we have missed some publications, or that some trials were not published. Estimation of the extent of this bias is difficult.

The attraction of therapy with megadoses of vitamins is easy to understand. It would be wonderful if a cheap therapy with few side effects could be found for diseases for which existing therapies are not fully satisfactory. However, besides the side effects mentioned, several serious side effects of (very) high dosages have been reported: pyridoxine may cause sensory neuropathy and nicotinic acid may induce severe liver dysfunction (Schaumburg et al 1983; Goldstein 1988).

Encouraging results have been reported only in autism. In most trials, vitamin B6 was combined with magnesium. Also, the dosis of vitamin B6 in these studies was approximately three times higher than in the other studies. Apart from pathophysiological differences, these points might explain why positive results were reported only here. It must also be emphasized that the evidence is limited, and that the results must be confirmed

in further, methodologically better trials. For all other indications there seems to be no adequate support from controlled trials to justify supplementation with either niacin, vitamin B6, or combinations of vitamins.

This work was supported by a grant of the Dutch Ministry of Welfare, Public Health and Cultural Affairs (Project 87-35).

References

- Altman H, Mehta D, Evenson RC, Sletten IW (1973): Behavioral effects of drug therapy on psychogeriatric inpatients. II. Multivitamin supplement. *J Am Geriatr Soc* 21:249-252.
- Ananth JV, Ban TA, Lehmann HE, Bennett J (1970): Nicotinic acid in the prevention and treatment of methionine-induced exacerbation of psychopathology in schizophrenics. *Can Psychiatr Assoc J* 15:15-20.
- Ananth JV, Vacaflor L, Kekhwa G, Sterlin C, Ban TA (1972): Nicotinic acid in the treatment of newly admitted schizophrenic patients: A placebo-controlled study. *Int J Clin Pharmacol Ther Toxicol* 5:406-410.
- Ananth JV, Ban TA, Lehmann HE (1973): Potentiation of therapeutic effects of nicotinic acid by pyridoxine in chronic schizophrenics. *Can Med Assoc J* 18:377-383.
- Anonymous (1987): Nomenclature policy: Generic descriptors and trivial names for vitamins and related compounds. *J Nutr* 117:7-14.
- Arnold LE, Christopher J, Huestis RD, Smeitzer DJ (1978): Megavitamins for minimal brain dysfunction. A placebo-controlled study. *JAMA* 240:2642-2643.
- Ashby WR, Collins GH, Basset M (1960): The effects of nicotinic acid, nicotinamide, and placebo on the chronic schizophrenic. *J Ment Sci* 106:1555-1559.
- Ban TA, Lehmann HE, Deutsch M (1977): Negative findings with megavitamins in schizophrenic patients. Preliminary report. *Commun Psychopharmacol* 1:119-122.
- Barthelemy C, Garreau B, Leddet I, et al (1980): Effets cliniques et biologiques de l'administration orale du magnésium associé à la vitamine B6 sur certains troubles observés dans l'autisme infantile. *Thérapie* 35:627-632.
- Bennett FC, McClelland S, Kriegsmann EA, Andrus LB, Sells CJ (1983): Vitamin and mineral supplementation in Down's syndrome. *Pediatrics* 72:707-713.
- Benton D, Buts JP (1990): Vitamin/mineral supplementation and intelligence. *Lancet* 335:1158-1160.
- Benton D, Roberts G (1988): Effect of vitamin and mineral supplementation on intelligence of a sample of schoolchildren. *Lancet* i:140-143.
- Bidder RT, Gary P, Newcombe RG, Evans BK, Hughes M (1989): The effects of multivitamins and minerals on children with Down's syndrome. *Dev Med Child Neurol* 31:532-537.
- Blass JP, Gleason P, Brush D, DiPonte P, Thaler H (1988): Thiamine and Alzheimer's disease. *Arch Neurol* 45:833-835.
- Boggs UR, Scheaf A, Santoro D, Ritzman R (1985): The effect of nutrient supplements on the biological and psychological characteristics of low IQ, preschool children. *J Orthomol Psychiatry* 14:97-127.
- Braunwald E, Isselbacher KJ, Petersdorf RG, Wilson JD, Martin JB, Fauci AS (eds) (1987): *Harrison's Principles of Internal Medicine*. New York: McGraw-Hill.
- Bumbalo TS, Morelewicz HV, Berens DL (1964): Treatment of Down's syndrome with the "U" series of drugs. *JAMA* 187:361.
- Chanowitz J, Ellman G, Silverstein CI, Zingarelli G, Ganger E (1985): Thyroid and vitamin-mineral supplement fail to improve IQ of mentally retarded adults. *Am J Ment Defic* 90:217-219.

- Chomé J, Paul T, Pudel V (1984): Testpsychologische Untersuchung bei älteren Menschen mit subklinischem Vitaminmangel. *Ernährungs Umschau* 31:12-16.
- Coburn SP, Schaltenbrand WE, Mahuren JD, Clausman RJ, Townsend D (1983): Effect of megavitamin treatment on mental performance and plasma vitamin B6 concentrations in mentally retarded young adults. *Am J Clin Nutr* 38:352-355.
- Coleman M, Steinberg G, Tippet J, et al (1979): A preliminary study of pyridoxine administration in a subgroup of hyperkinetic children: A double-blind crossover comparison with methylphenidate. *Biol Psychiatry* 14:741-751.
- Coleman M, Sobel S, Bhagavan HN, et al (1985): A double-blind study of vitamin B6 in Down's syndrome infants. Part 1—Clinical and biochemical results. *J Ment Defic Res* 29:233-240. [Part 2: Frager et al (1985)]
- Colgan M, Colgan L (1984): Do nutrient supplements and dietary changes affect learning and emotional reactions of children with learning difficulties? A controlled series of 16 cases. *Nutr Health* 3:69-77.
- Crombie IK, Todman J, McNeill G, du V Florey C, Menzies I, Kennedy RA (1990): Effect of vitamin and mineral supplementation on verbal and non-verbal reasoning of schoolchildren. *Lancet* 335:744-747.
- Denson R (1962): Nicotinamide in the treatment of schizophrenia. *Dis Nerv System* 23:167-172.
- Deutsch M, Ananth JV, Ban TA (1977): Nicotinic acid in the treatment of chronic hospitalized schizophrenic patients: A placebo controlled clinical study. *Psychopharmacol Bull* 13:21-23.
- Dries WJH, Schreurs WHP, Buyze G, Schakelaar AJ (1978): Onderzoek naar de invloed van vitaminesuppletie op het gedrag van psychiatrische patiënten. *Tijdschrift voor Psychiatrie* 20:307-314.
- Ellis NR, Tomporowski PD (1983): Vitamin/mineral supplements and intelligence of institutionalized mentally retarded adults. *Am J Ment Defic* 88:211-214.
- Ellman G, Silverstein CI, Zingarelli G, Schafer EWP, Silverstein L (1984): Vitamin-mineral supplement fails to improve IQ of mentally retarded young adults. *Am J Ment Defic* 88:688-691.
- Frager J, Barnet A, Weiss I, Coleman M (1985): A double-blind study of vitamin B6 in Down's syndrome infants. Part 2—Cortical auditory-evoked potentials. *J Ment Defic Res* 29:241-246.
- Friedrich W (ed) (1988): *Vitamins*. Berlin: de Gruyter.
- Goldstein MR (1988): Potential problems with the widespread use of niacin. *Am J Med* 85:881.
- Greenbaum GHC (1970): An evaluation of niacinamide in the treatment of childhood schizophrenia. *Am J Psychiatry* 127:129-132.
- Harrell RF, Capp RH, Davis DR, Peerless J, Ravitz LR (1981): Can nutritional supplements help mentally retarded children? An exploratory study. *Proc Natl Acad Sci USA* 78:574-578.
- Haslam RHA, Dalby JT, Rademaker AW (1984): Effects of megavitamin therapy on children with attention deficit disorders. *Pediatrics* 74:103-111.
- Heseker H, Kübler W, Westenhöfer J, Pudel V (1990): Psychische Veränderungen als Frühzeichen einer suboptimalen Vitaminversorgung. *Ernährungs Umschau* 37:87-94.
- Hodges RE (ed) (1980): *Nutrition in Medical Practice*. Philadelphia: WB Saunders.
- Hoffer A (1962): *Niacin Therapy in Psychiatry*. Springfield: Charles C Thomas.
- Hoffer A, Osmond H, Callbeck MJ, Kahan I (1957): Treatment of schizophrenia with nicotinic acid and nicotinamide. *J Clin Exp Psychopathol Q Rev Psychiatry Neurol* 18:131-158.
- Jonas C, Étienne T, Barthelemy C, Jouve J, Mariotte N (1984): Intérêt clinique et biochimique de l'association vitamine B6 + magnésium dans le traitement de l'autisme résiduel à l'âge adulte. *Thérapie* 39:661-669.
- Kerschner J, Grekin R, Hawke WA, Cutler P (1977): Pilot study of high-protein, high-vitamin, low-carbohydrate, sugar-free diet in learning-disabled children. *Can Med Assoc J* 117:212.

- Lipton MA, Mailman RB, Nemeroff CB (1987): Vitamins, megavitamin therapy, and the nervous system. In: Wurtman RJ, Wurtman JJ (eds), *Nutrition and the Brain*, vol 3. New York: Raven Press.
- Loriaux SM, Deijen JB, Orlebeke JF, De Swart JH (1985): The effects of nicotinic acid and xanthinol nicotinate on human memory in different categories of age. A double-blind study. *Psychopharmacology* 87:390-395.
- Lu L, Stotsky BA, Cole JO (1971): A controlled study of drugs in long-term geriatric psychiatric patients. A double-blind comparison of pentylenetetrazol, papaverine, and niacin. *Arch Gen Psychiatry* 25:284-288.
- Martineau J, Barthelemy C, Garreau B, Lelord G (1985): Vitamin B6, magnesium, and combined B6-Mg: Therapeutic effects in childhood autism. *Biol Psychiatry* 20:467-478.
- Maseck D (1980): Vitamins: The get-smart pills? *J Orthomol Psychiatry* 9:58-65.
- McGrath SD, O'Brien PF, Power PJ, Shea JR (1972): Nicotinamide treatment of schizophrenia. *Schizophr Bull* 5:74-76.
- Menolascino FJ, Donaldson JY, Gallagher TH, et al (1989): Vitamin supplements and purported learning enhancement in mentally retarded children. *J Nutr Sci Vitaminol* 35:181-192.
- Molimard R, Marillaud A, Paille A, Le Devehat C, Lemoine A, Dougny M (1980): Impairment of memorization by high doses of pyridoxine in man. *Biomedicine* 32:88-92.
- Naismith DJ, Nelson M, Burley VJ, Gatenby SJ (1988): Can children's intelligence be increased by vitamin and mineral supplements? *Lancet* ii:335.
- Nestoros JN, Ban TA, Lehmann HE (1977): Transmethylation hypothesis of schizophrenia: Methionine and nicotinic acid. *Int Pharmacopsychiatry* 12:215-246.
- O'Reilly PO (1955): Nicotinic acid therapy and the chronic schizophrenic. A pilot study. *Dis Nerv Syst* 16:67-72.
- Pueschel SM, Reed RB, Cronk CE, Goldstein BI (1980): 5-Hydroxytryptophan and pyridoxine. *Am J Dis Child* 134:838-844.
- Ramsay RA, Ban TA, Lehmann HE, Saxena BM, Bennett J (1970): Nicotinic acid as adjuvant therapy in newly admitted schizophrenic patients. *Can Med Assoc J* 102:939-942.
- Rimland B, Callaway E, Dreyfus P (1978): The effect of high doses of vitamin B6 on autistic children: A double-blind crossover study. *Am J Psychiatry* 135:472-475.
- Schaumburg H, Kaplan J, Windebank A, et al (1983): Sensory neuropathy from pyridoxine abuse: A new megavitamin syndrome. *N Engl J Med* 309:445-448.
- Sehdev HS, Olson JL (1974): Nicotinic acid therapy in chronic schizophrenia. *Compr Psychiatry* 15:511-517.
- Smith GF, Spiker D, Peterson CP, Cichetti D, Justine P (1984): Use of megadoses of vitamins with minerals in Down's syndrome. *J Pediatr* 105:228-234.
- Vallely JF, Lovegrove TD, Hobbs GE (1971): Nicotinic acid and nicotinamide in the treatment of chronic schizophrenia. *Can Psychiatr Assoc J* 16:433-435.
- Weathers C (1983): Effects of nutritional supplementation on IQ and certain other variables associated with Down's syndrome. *Am J Ment Defic* 88:214-217.
- Wittenborn JR, Weber ESP, Brown M (1973): Niacin in the long-term treatment of schizophrenia. *Arch Gen Psychiatry* 28:308-315.

Vitamin B6 in the treatment of the premenstrual syndrome—a review

JOS KLEIJNEN, GERBEN TER RIET, PAUL KNIPSCHILD

Summary. A search of the literature yielded 12 controlled trials on vitamin B6 in the treatment of the premenstrual syndrome. These are discussed with emphasis on methodological aspects. A major drawback of the trials is the limited number of patients included. The existing evidence of positive effects of vitamin B6 is weak, and some well-designed trials with positive results would be needed to change this view.

Many theories on the aetiology of the premenstrual syndrome (PMS) have been proposed, including speculation about the effects of hormonal imbalances involving oestrogen, progesterone and prolactin, and about neuroendocrine dysfunction. Dysfunction of the pituitary gland production of endogenous opiates and/or α -melanocyte-stimulating hormone might explain most of the symptoms (Reid & Yen 1981). An alternative explanation might be decreased sensitivity of progesterone receptors in several organs. In any event, no generally accepted explanation of the aetiology currently exists.

Various therapies have been advocated in the treatment of PMS. One of them is the use of megadoses of vitamin B6 (pyridoxine). Pyridoxal 5' phosphate, the active form of vitamin B6, serves as the coenzyme of a wide variety of enzymes of amino acid metabolism (Friedrich 1988). It serves for instance as a cofactor in the metabolism of tryptophan (the precursor of serotonin), and also in the metabolism of tyrosine (leading to dopamine and noradrenalin) and glutamate (leading to gamma-aminobutyric acid). Low levels of dopamine and serotonin

lead to high levels of prolactin and aldosterone, explaining the fluid retention. Against this background, postulation of beneficial effects of vitamin B6 supplementation on both physical (fluid retention) and psychological (effect on neurotransmitters) symptoms related to PMS seems reasonable.

However, implementation in regular practice can only be considered if clinical practice shows positive results. In a survey of 282 randomly chosen general practitioners in the Netherlands, 36% reported that they believed in the efficacy of vitamin B6 in PMS (Knipschild *et al.* 1990). The best way to show its efficacy is, of course, to perform randomized, placebo-controlled trials. The trials found after searching the literature are reviewed here.

Materials and methods

Identification and selection of relevant studies

Trials were only eligible if an index and a control group were included and if vitamin B6 was given to one of the treated groups. Controlled studies in animal models were excluded. Experiments were found by screening Index Medicus from 1966 to 1989 (keywords: pyridoxine-administration-and-dosage and pyridoxine-therapeutic-use), and by screening Excerpta Medica, section 10 (Obstetrics and Gynecology) from 1966-1989 (keywords: pyridoxine and vitamin B6). Further experiments were found by checking references extensively and by personal communication with authors of experiments. We also visited Hoff-

Department of Epidemiology/Health Care Research,
University of Limburg, Postbox 616, 6200 MD
Maastricht, The Netherlands

J. KLEIJNEN

G. TER RIET

P. KNIPSCHILD *Professor*

Correspondence to: J. Kleijnen

man-La Roche, in Basel, in an attempt to ascertain unpublished studies.

We found 10 randomized double-blind trials on the effects of vitamin B6 in patients with PMS. In two other double-blind trials the effect of vitamin B6 on premenstrual mastalgia was studied. These 12 studies are listed in Table 1 and are considered in detail in Tables 2 and 3.

One study (Day 1979) was excluded because it is not really a cross-over study (no random allocation; all patients received vitamin B6 during cycle 2, 3 and 4 and placebo during cycle 1, 5 and 6). Another study was excluded because only one woman was treated (Mattes & Martin 1982). These two studies both suggested positive results of vitamin B6 treatment.

Assessment of methodological quality of published studies

General criteria such as comparability of the treatment groups (prestratification, randomization), the interventions (adequate placebos) and the effect measurements (double blinding) should be met in every therapeutic trial. Special attention must be paid to some specific aspects of trials on PMS. The predominant symptoms of PMS are abdominal bloating, breast pain or tenderness, headache, oedema, fatigue, depression and irritability, starting somewhere in the second half of the cycle, with relief at or soon after the onset of menstruation. A clear definition of PMS does not exist, thus authors of trials should indicate what their inclusion criteria were and which symptoms were assessed in their investigation. An overall, global index of the complaints appears appropriate, since the impact of individual symptoms on the total discomfort may vary from patient to patient. An adequate description of the way in which these symptoms have been assessed (which questionnaire on depression, which scales for the assessment of pain etc.) is also important.

Furthermore, it is advisable to extend the investigation over several cycles, especially in experiments with a limited number of patients, since there may be considerable variability of the severity of the symptoms both within patients in different cycles and between patients.

To explore the possibility that an increasing likelihood of bias (with an increasing number of shortcomings) is reflected in the results of the trials, we used the following criteria for a methodological assessment of the experiments:

(A) Characteristics of the women adequately described (description of symptoms, duration, severity, and previous treatments). (B) At least 50 women in each group. (C) Prestratification (matching) on relevant prognostic factors. (D) Randomization. (E) Check of prognostic comparability. (F) Subjects blinded. (G) Assessors blinded. (H) Effect measurements well described (reproducible). (I) Use of other concurrent treatments or vitamins has been checked and (J) presentation of the results in such a manner that the analysis can be checked by the reader.

It could be argued that other criteria should be used for the methodological assessment. However, we have selected well established methodological criteria, as have been described by Meinert (1986), and our assessment can be checked by the reader (Table 2).

Results

The main characteristics and results of the studies are summarized in Table 1.

The dosages which were used varied from 50 mg to 500 mg pyridoxine a day (the recommended daily allowance in healthy adults is 2 mg/day). No relation to the results of the experiments can be seen regarding different dosage levels. The duration of the experiments varied from 1 to 12 cycles; most experiments lasted 2 or 3 cycles. Individual symptoms, or scales assessing many symptoms, e.g., 47 in a scale developed by Moos (1969), were mostly used for the measurement of the treatment effects. These were in general reasonably well described. However, the number of women, mostly 20–40 in each group, was limited. The only exception is the multi-centre study reported by Williams *et al.* (1985) (200 patients in each group). Overall treatment effects additional to the assessment of individual symptoms were measured in only three trials (Barr 1984, Hagen *et al.* 1985 and Williams *et al.* 1985). Only trivial or no side effects were reported in the studies presented.

Abraham & Hargrove (1980), Barr (1984) and Hallman (1987) reported positive effects, but generally the symptoms did not disappear completely. Stokes & Mendels (1972), Williams *et al.* (1985), Smallwood *et al.* (1986), Kendall & Schnurr (1987), and Doll *et al.* (1989) found ambiguous effects; overall the results of these studies were disappointing. Colin (1982), Hagen

Table 1. Characteristics of controlled trials of vitamin B6 in the treatment of the premenstrual syndrome

Study	Entry criteria	Number and design	Daily dosage (mg)	No. of cycles,* effect measures	Outcome data
<i>Positive result</i>					
Abraham & Hargrove (1980)	Moderate to severe symptoms, no medication and no OC	25 Cross-over	500	2* 3, 19 symptoms	Significant B-6 effect in excess of placebo effect in 21 subjects
Barr (1984)	— (most patients were not on any OC)	36 Cross-over	100	2* 2, 8 symptoms	Improvement for 83% on B-6 and 28% on placebo
Hallman & Orelund (1987)	Severe symptoms, no hormonal therapy	32 Cross-over	300	2* 4, 14 symptoms	Significant improvements found for 8 symptoms
<i>Ambiguous result</i>					
Stokes & Mendels (1972)	Premenstrual tension, depression	13 Cross-over	50	8-12, 47 symptoms	Not clear, 'disappointing'
Williams <i>et al.</i> (1965)	Medication and OC were allowed	204/230 Parallel	100	3, 9 symptoms	Improvement for 82% on B-6 and 70% on placebo
Smallwood <i>et al.</i> (1986)	Premenstrual mastalgia, paracetamol was allowed	42, Cross-over	200	2* 2, pain, tenderness	No significant improvement; positive trend for tenderness on B-6
Kendall & Schnurr (1987)	Moderate to severe symptoms, no medication and no OC	29/26 Parallel	150	2, 47 symptoms	Significant amount of symptomatology remained (positive for behaviour, dizziness and nausea)
Doll <i>et al.</i> (1989)	Moderate to severe symptoms, no medication, OC were allowed	32, Cross-over	50	2* 3, 9 symptoms	Significant beneficial effect on emotional symptoms, no effect on other symptoms
<i>Negative result</i>					
Colin (1982)	— Mastalgia	17/15 Parallel, random?	500	2-4 months, pain, temperature	Improvement for 59% on B-6 and 60% on placebo
Hagen <i>et al.</i> (1985)	No medication and no OC	34, Cross-over	100	2* 2, scale of 6 symptoms	Improvement for 4% on B-6 and 21% on placebo using visual analogue scale
Malmgren <i>et al.</i> (1987)	No medication and no OC	19, Cross-over, random?	300	2* 1, many symptoms	No significant improvement of symptoms
Van den Berg <i>et al.</i> (1989)	No medication and no OC	34/42 Parallel	120	4, 47 symptoms	No difference

* 2* K = two periods of K cycles (cross-over).
OC = oral contraceptives.

et al. (1985), Malmgren *et al.* (1987), and Van den Berg *et al.* (1989) could not find any positive effects.

Table 2 shows that all trials had at least two methodological shortcomings. The mean 'score'

for the trials with positive, ambiguous and negative results is 4.5, 6.2 and 4.1 respectively (maximum is 10), indicating that there is no trend in any direction according to the likelihood of bias (number of methodological shortcomings).

Table 2. Methodological score assigned to the 12 trials

Study	Methodological criteria										TOT
	A	B	C	D	E	F	G	H	I	J	
Positive result											
Abraham & Hargrove (1980)	±	-	-	+	+	+	+	+	-	+	6.5
Barr (1984)	-	-	-	+	-	+	+	-	-	-	3
Hallman & Öreland (1987)	±	-	-	+	-	+	+	±	-	-	4
Ambiguous result											
Stokes & Mendels (1972)	-	-	-	+	-	+	+	-	-	-	3
Williams <i>et al.</i> (1985)	±	+	±	+	-	+	+	+	+	+	8
Smallwood <i>et al.</i> (1986)	±	-	-	+	+	+	+	-	+	+	6.5
Kendall & Schnurr (1987)	±	-	+	+	+	+	+	+	-	±	7
Doll <i>et al.</i> (1989)	+	-	-	+	±	+	+	+	-	+	6.5
Negative result											
Colin (1982)	-	-	-	±	-	+	+	-	-	-	2.5
Hagen <i>et al.</i> (1985)	±	-	-	+	-	+	+	+	-	±	5
Malmgren <i>et al.</i> (1987)	+	-	-	±	-	+	+	-	-	-	3.5
Van den Berg <i>et al.</i> (1989)	±	-	-	+	+	+	+	+	-	-	5.5

Explanation of scores: +, item is scored; -, item is not scored; ±, item is partly scored; TOT, number of items which were scored.

A, Characteristics of subjects adequately described (description of symptoms, duration, severity, and previous treatments); B, at least 50 women per group; C, prestratification (matching) on relevant prognostic factors; D, randomization; E, check of prognostic comparability; F, subjects blinded; G, assessor blinded; H, effect measurement well described (reproducible); I, use of other concurrent treatments or vitamins has been checked; J, presentation of the results in such a manner that the analysis can be checked by the reader.

Table 3 shows which symptoms improved and which symptoms were not influenced by vitamin B6 treatment in the trials reporting on individual symptoms. The variation between trials turns out to be large and no trends are visible.

Discussion

Although PMS is prevalent among women of child bearing age, the number of participants in most studies was disappointingly low. Furthermore, compliance, blinding and prognostic comparability was checked in only 3, 0 and 1 trials respectively.

It could be argued that in cross-over studies fewer subjects need to be admitted because each subject serves as her own control. In these designs, however, the basic assumption that the women are prognostically comparable at the cross-over to the alternative treatment may not be met (inadequate washout, carry-over effects). In the eight cross-over studies considered in this review only Hallman & Öreland (1987) and Hagen *et al.* (1985) included wash-

out periods, unless one assumes that the post menstrual phase of 2 weeks is sufficient to assure prognostic comparability of the groups at the beginning of the second treatment phase. Estimation of the bias introduced by carry-over effects is very difficult. Interestingly of the four studies with a parallel group design, two showed negative results and two ambiguous results. All positive assessments resulted from cross-over studies.

We did not pool the results of the trials. The methodological quality must obviously be high before statistical pooling makes sense. Moreover, there were differences in the characteristics of the women, the dosage of vitamin B6, the duration of the intervention and the outcome measurements.

Publication bias is important in every review article. Trials with negative results are less likely to be submitted and published than trials with positive results, especially if limited numbers of subjects have participated. This might well be the case for trials of vitamin B6 in PMS. The magnitude of this bias is difficult to estimate.

Table 3. The effects of vitamin B6 on symptoms of the premenstrual syndrome

Study	Positive effects	No difference
Hagen <i>et al.</i> (1985)		Depression, irritability, bloatedness, weight gain, headache, breast tenderness, fatigue
Williams <i>et al.</i> (1985)		Depression, irritability, swelling/bloating, headache, breast tenderness, tension, violent feelings, coordination
Smallwood <i>et al.</i> (1986)	Breast tenderness	Breast pain
Kendall & Schnurr (1987)	Dizziness, nausea, poor performance, social activities	Many including: depression, anxiety
Hallmann & Orelund (1987)	Depression, irritability, swollenness, headache, hyperactivity, clumsiness, skin/hair changes	Tension, anxiety, apathy, sleeplessness, tiredness, appetite
Van den Berg <i>et al.</i> (1989)		Many including: depression, irritability, bloating, weight gain, headache, mastalgia, tension, anxiety
Doll <i>et al.</i> (1989)	Depression, irritability, tiredness	Swollenness, headache, breast discomfort, stomach cramps, backache

No positive or ambiguous effects of vitamin B6 have been observed in controlled trials on related subjects, such as the treatment of depression related to the use of oral contraceptives (Adams *et al.* 1973, Adams *et al.* 1974), inhibition of lactation (Foukas 1972, Foukas 1973, Macdonald *et al.* 1976, Canales *et al.* 1976), and in nausea and vomiting of pregnancy (Hesseltine 1946, General Practitioner Research Group 1963, Hillman *et al.* 1963).

In future studies, with larger groups of women, more attention must be paid to assuring prognostic comparability of the treatment groups, and to checking of compliance and blinding. A question remains about which dose of the vitamin should be used, what the optimal duration of treatment is, and whether perhaps only special subgroups (e.g., with relative deficiencies of vitamin B6) might benefit from vitamin B6 supplementation. At the moment there is no evidence that vitamin B6 is efficacious in the treatment of patients with PMS.

Acknowledgments

This work was supported by a grant from the Dutch Ministry of Welfare, Public Health and Cultural Affairs.

References

- Abraham G. E. & Hargrove J. T. (1980) Effect of vitamin B6 on premenstrual tension syndromes: a double blind crossover study. *Infertility* 3, 155-165.
- Adams P. W., Wynn V., Rose D. P., Seed M., Folkard J. & Strong R. (1973) Effect of pyridoxine hydrochloride (vitamin B6) upon depression associated with oral contraception. *Lancet* i, 897-904.
- Adams P. W., Wynn V., Seed M. & Folkard J. (1974) Vitamin B6, depression, and oral contraception. *Lancet* ii, 516-517.
- Barr W. (1984) Pyridoxine supplements in the premenstrual syndrome. *Practitioner* 228, 425-427.
- Canales E. S., Soria J., Zárate A., Mason M. & Molina M. (1976) The influence of pyridoxine on prolactin secretion and milk production in women. *Br J Obstet Gynaecol* 83, 387-388.
- Colin C. (1982) Etudes contrôlées de l'administration orale de progestagènes, d'un antioestrogène et de vitamine B6 dans le traitement des mastodynies. *Rev Méd Brux* 3, 605-609.
- Day J. B. (1979) Clinical trials in the premenstrual syndrome. *Curr Med Res Opin* 6, suppl. 5, 40-45.
- Doll H., Brown S., Thurston A. & Vessey M. (1989) Pyridoxine (vitamin B6) and the premenstrual syndrome: a randomized crossover trial. *J R Coll Gen Pract* 39, 364-368.
- Foukas M. (1972) Laktationshemmende Wirkung des Pyridoxins. *Dtsch Med Wochenschr* 97, 396-397.

852 J. Kleijnen et al.

- Foukas M. D. (1973) An antilactogenic effect of pyridoxine. *J Obstet Gynaecol Br Commonwealth* **80**, 718-720.
- Friedrich W. (1988) Vitamin B6. In *Vitamins*, de Gruyter, Berlin, New York, pp 543-618.
- General Practitioner Research Group (1963) Meclozine and pyridoxine in pregnancy sickness. *Practitioner* **190**, 251-253.
- Hägen I., Nesheim B-I. & Tuntland T. (1985) No effect of vitamin B6 against premenstrual tension. A controlled clinical study. *Acta Obstet Gynecol Scand* **64**, 667-670.
- Hallman J. & Orelund L. (1987) Therapeutic effect of vitamin B6 in the treatment of premenstrual syndrome. A double-blind cross-over study. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* **88**, 1-15.
- Hesseltine H. C. (1946) Pyridoxine failure in nausea and vomiting of pregnancy. *Am J Obstet Gynecol* **51**, 82-86.
- Hillman R. W., Cabaud P. G., Nilsson D. E., Arpin P. D. & Tufano R. J. (1963) Pyridoxine supplementation during pregnancy. *Am J Clin Nutr* **12**, 427-430.
- Kendall K. E. & Schnurr P. P. (1987) The effects of vitamin B6 supplementation on premenstrual symptoms. *Obstet Gynecol* **70**, 145-149.
- Knipschild P., Kleijnen J. & ter Riet G. (1990) Belief in the efficacy of alternative medicine among general practitioners in the Netherlands. *Soc Sci Med*, in press.
- Macdonald H. N., Collins Y. D., Tobin M. J. W. & Wijayaratne D. N. (1976) The failure of pyridoxine in suppression of puerperal lactation. *Br J Obstet Gynaecol* **83**, 54-55.
- Malmgren R., Collins A. & Nilsson C-G. (1987) Platelet serotonin uptake and effects of vitamin B6-treatment in premenstrual tension. *Neuropsychobiology* **18**, 83-88.
- Mattes J. A. & Martin D. (1982) Pyridoxine in premenstrual depression. *Hum Nutr Appl Nutr* **36A**, 131-133.
- Meinert C. L. (1986) *Clinical Trials. Design, Conduct and Analysis*. Oxford University Press, New York; Oxford, pp 65-70.
- Moos R. H. (1969) Typology of menstrual cycle symptoms. *Am J Obstet Gynecol* **103**, 390-402.
- Reid R. L. & Yen S. S. C. (1981) Premenstrual syndrome. *Am J Obstet Gynecol* **139**, 85-104.
- Smallwood J., Ah-Kye D. & Taylor I. (1986) Vitamin B6 in the treatment of pre-menstrual mastalgia. *Br J Clin Pract* **40**, 532-533.
- Stokes J. & Mendels J. (1972) Pyridoxine and premenstrual tension. *Lancet* **i**, 1177-1178.
- Van den Berg H., Schrijver J., Bruinse H. W. & Van der Ploeg H. M. (1989) Vitamin B6 and premenstrual syndrome (PMS). *Voeding* **50**, 58-62.
- Williams M. J., Harris R. I. & Dean B. C. (1985) Controlled trial of pyridoxine in the premenstrual syndrome. *J Int Med Res* **13**, 174-179.

Received 2 October 1989

Accepted 14 May 1990

CHAPTER 8

EVENING PRIMROSE OIL*

Jos Kleijnen, Gerben ter Riet, Paul Knipschild

* Published in Pharm Weekbl 1989;124:418-423

Summary

The oil from the seeds of the evening primrose (*Oenothera biennis*) is used as a medicine in many diseases. The positive effects of evening primrose oil are often attributed to one of its components (gammalinolenic acid). A review is given of controlled clinical trials examining the effects of evening primrose oil for various indications. Most studies reviewed show many methodological defects. For several disorders further studies on the effectiveness of evening primrose oil and/or gammalinolenic acid seem warranted. As long as its effectiveness has not been proven we will not recommend the use of evening primrose oil.

Introduction

The oil that can be pressed out of the seeds of the evening primrose (*Oenothera biennis*) is marketed under several trade names, such as Efamol and Naudicelle, and is administered in many disorders. Their capsules contain about 72% of linoleic acid, 9% of gammalinolenic acid (GLA), and 10IE of vitamin E as an antioxidant. The positive results obtained from groups of patients with all kinds of disorders are often attributed to GLA. It has been established that GLA is much more active, biologically, than linoleic acid in, for example, the correction of symptoms related with linoleic acid deficiencies. As, generally, relatively large quantities of these expensive and sizable capsules must be swallowed during a long period, the question of its effectiveness arises. In this article a review is given of the state-of-the-art, based on controlled trials in human beings.

Mechanisms

Polyunsaturated fatty acids such as linoleic acid and its metabolite GLA are precursors of important active substances such as prostaglandins (PG), thromboxans, and leukotrienes (figure 1). These substances play a role in many physiological and pathological processes.¹

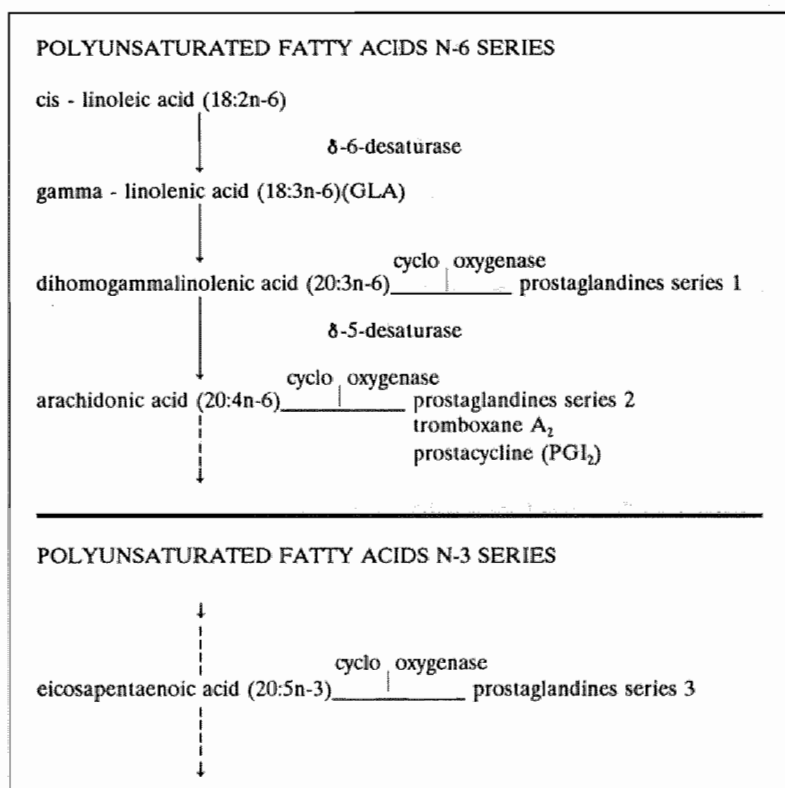


Figure 1. Formation of prostaglandines and related compounds from linoleic acid.

Although much is known by now about the formation and the effects of these substances, much remains to be elucidated about different factors that influence these processes. Examples of such factors are the effects in the administration of GLA, vitamins, minerals, alcohol, drugs such as acetylsalicylic acid, and mechanisms that play a role in all kinds of diseases. A balance seems to exist in the formation of three series of prostaglandins in which fatty acids compete for the enzyme cyclo-oxygenase. The quantity of substrate that is available, would play a role in this process.

An important theory states that under various circumstances a (relative) GLA deficiency is caused by a lack of or an inhibition of the enzyme delta-6-desaturase (figure 1).^{2,3} As a result of this, relatively more arachidonic acid would be available to form the PG2 series by the effect of cyclo-oxygenase. This can occur because arachidonic acid can be obtained, besides through formation from dihomogamma-linolenic acid, through nutrition (meat, eggs) as well. By administering GLA, the first slow step (figure 1) is by-passed and the relationship between GLA and arachidonic acid changes. In that case the formation of the PG1 series would be affected positively. Particularly in inflammations, this PG1 series gives less acute reactions than the PG2 series that develops from arachidonic acid. Eicosapentaenoic acid (fish consumption) is converted, by the effect of cyclo-oxygenase, into the PG3 series that is also less active than the PG2 series.

Controlled studies

Trials were found by consulting a bibliography (by C. Toplack) obtained via the Efamol Research Institute, by checking of references, and a MEDLINE search of studies encompassing the period 1963 to 1988. The following combination of key words was used: linolenic acids and clinical trials, linolenic acids and (premenstrual syndrome or premenstrual tension or Raynaud's disease or diabetes mellitus or eczema or dermatitis, atopic or Sjögren's syndrome), and finally fatty acids, essential and (premenstrual syndrome or etc....).

Disorders in which evening primrose oil is credited for a positive effect are, among others, multiple sclerosis, rheumatoid arthritis and other disorders related to this disease, atopic dermatitis, premenstrual syndrome and mastalgia, cancer, Raynaud's phenomenon and cardiovascular disorders. Often, in cardiovascular disorders, the effects on risk factors only are described.³ It strikes one that it concerns diseases that sometimes are hard to examine, because they show many subjective complaints and a whimsical, spontaneous course. When reading the trials in humans critically, it strikes one that there is much fault to find with the methodology used. The design features, the intervention and the assessment of the effects of the trials are often briefly described, and the number of participants used is often limited. In controlled trials in humans features such as prognostically comparable groups, double blinding, and a clear description of the set-up, execution and effect assessment are, in our opinion, determinants for their quality and value. Particularly relatively better trials are quoted here. The rest of the trials with control groups have been entered in the list of references.⁴⁻²⁰ As evening primrose oil Efamol was used in the undermentioned trials, with the exception of Bates et al. (Naudicelle).

Multiple sclerosis

In 1977 and 1978 Bates et al. published two trials (both randomized and double blind) on the treatment of multiple sclerosis with polyunsaturated fatty acids. In the first trial the effect of 8 capsules of evening primrose oil daily was tested during two years in patients with chronic, progressive complaints. The control group received capsules containing oleic acid. In each group there were 38 patients. The second trial had the same

design, but these patients had acute complaints and remissions (29 patients in each group). The two trials did not show differences between the evening primrose oil group and the control group. This trial was criticised for the fact that the capsules contained a colouring agent that inhibited the conversion of GLA into prostaglandins. Moreover, according to the critics, the dosage of GLA of 8 capsules daily (which is equal to 3.4g of linoleic acid and 360mg of GLA) was too low.²¹⁻²³

Trials with administration of linoleic acid only showed varying results with a trend towards positive results, as the illness was in an earlier stage.²⁴ Taking the line that an early diagnosis can sometimes give problems and that the follow-up for this disorder was rather short (maximally two years) reliable results can only be obtained if larger groups are studied over an even longer period of time. Therefore, for the time being, no definitive conclusions regarding the therapy with linoleic acid and GLA of multiple sclerosis can be drawn.

Rheumatoid arthritis

Studies have been undertaken for years now on the effect of nutritional supplements and diets in this disorder, so far without any striking results.

Belch et al. examined in 49 patients with rheumatoid arthritis in a randomized, double-blind trial whether evening primrose oil, in combination with fish oil or not, had any effect on the use of NSAID's, on the duration of the morning stiffness, on the pinching power of the two hands, on the condition of the joints, and on pain on a 10 cm visual analogue scale and finally whether the patients benefited from the treatment.²⁵ The dosage was 12 capsules of evening primrose oil daily (540mg of GLA) for 16 patients, 12 capsules of evening primrose oil and fish oil daily (450mg of GLA and 240mg of eicosapentaenic acid) for 15 patients, and 12 placebo capsules containing paraffine daily for 18 patients. All capsules contained vitamin E (120IE daily). During the first 3 months the capsules and the normal dosage of NSAID's were taken, after that period the patients were allowed to reduce their dosages of NSAID's. After 12 months all groups received a placebo without vitamin E during 3 months (single-blind). After one year 10 patients in the placebo group had withdrawn (aggravation of the symptoms) and 3 and 2 patients in the evening primrose oil-group and the evening primrose oil/fish oil-group respectively. In both treatment groups more patients could reduce their use of NSAID's or stop using them than in the placebo group. The clinical assessments did not show any differences, the subjective response was better for the index groups. After the 3 last months of the study (the placebo period) the condition of all the patients from the evening primrose oil-group and 80% of the patients from the combined group had deteriorated. As compared with 14% of the patients in the placebo group. The authors concluded that it is improbable that these fatty acids affect the prognosis or the course of the disease, but that they can be used in situations in which treatment with NSAID's is contra-indicated. We think that this last conclusion is premature at this moment.

Sjögren's syndrome

GLA could be useful as a precursor of PGE1 because this prostaglandin was able to prevent pathological manifestations in animal models of Sjögren's syndrome.²⁶ A confirmation of this theory in humans is not known to us. Sjögren's syndrome is characterized by a decrease of lachrymal and saliva production and a lympho/plasmacellular infiltrate in the glands. About half of the patients suffering from this disease also suffers from rheumatoid arthritis. Oxholm et al. examined in a double-blind cross-over trial the effect of administration of evening primrose oil, 6 capsules daily, for 8 weeks, or a placebo

(which was not described).²⁷ The allocation to the groups was random. In this orderly reported study the keratoconjunctivitis sicca was measured by means of 3 different tests, and the xerostomy (dry mouth) was measured by means of unstimulated sialometry. Assessments took place after 0, 4, 8, 12, and 16 weeks. There was no long-term follow-up. The subjective effects were not reported. The results of the group that received evening primrose oil were more positive for the three end points for keratoconjunctivitis sicca, but no positive effect was found on xerostomy. Raw data for these endpoints were not given. Unfortunately, only 24 patients were involved: 20 women and 4 men. We agree with the authors in their conclusion: "There is thus a need for further studies before the value of EFA (essential fatty acids) treatment of patients with primary SS (Sjögren's syndrome) can be finally stated."

Atopic dermatitis

Wright and Burton examined in a randomized, double-blind, cross-over study the effect of various dosages of evening primrose oil in 99 patients (60 adults and 39 children) with atopic dermatitis.²⁸ The patient continued the use of their usual medication, that is corticosteroids, anti-histamines and a neutral cream. The intervention took 12 weeks whereupon a cross-over took place. The capsules contained 360mg of linoleic acid and 45mg of GLA, the placebo consisted of capsules containing paraffine. The dosages were 4, 8 or 12 capsules daily for the adults and 2 or 4 capsules daily for the children. There were 16 patients in the group of adults and 3 patients in the group of children who had withdrawn. As end points a linear 10cm scale was used for scale, redness, itch and severity. Every three weeks there were assessments. In the analysis the subgroups were joined and the results are presented as differences between the scales used before and after the intervention. The usual medication was not described as endpoint and was not quantified. The results in the subgroups with the highest dosages showed a positive effect on all endpoints.

Bamford et al. examined in a randomized, double-blind, cross-over trial in 123 patients the effect of the same capsules Wright and Burton had used.²⁹ The dosage was 12 or 16 capsules daily for the adults and 4 or 8 capsules for the children. As a placebo capsules were used containing paraffine to which unlike the above mentioned trial also 10IE of vitamin E was added. Using 16 capsules daily this amounts to 160IE of vitamin E daily (that is about 16 times the officially recommended daily intake). The compliance was checked and described. Forty-six percent of the patients attained a compliance of 75%, and 65% of the testees attained a compliance of 50%. Thirty-one patients had withdrawn, of whom 14 during the treatment with evening primrose oil and 17 during treatment with a placebo. The endpoints were assessed on a 10-point scale for redness, scale, excoriation, lichenification and severity. Changes in the intake of regular medication were assessed, but their results cannot be found in the article. No differences were found between the placebo group and the therapy group, nor was there a dose-response-effect. As a possible explanation for the differences in the findings, compared with the findings by Wright and Burton, Bamford et al. point out that the patients in the trial of Wright and Burton had a worse clinical picture than their own patients, that there were differences in the initial arachidonic acid concentrations, and that another placebo without vitamin E was used. Other (brief) publications also show controversial results.^{5,13,16,18,20} In our view the effectiveness of evening primrose oil has not been demonstrated in this disorder.

Premenstrual syndrome (PMS) and mastalgia

For many women a satisfactory therapy for PMS cannot be found. Treatments with hormones, vitamins, anti-depressants, and diets often have no effect. Puolakka et al. studied the effect of 6 capsules of evening primrose oil daily in 30 women who had suffered a lot from complaints going with PMS since 9 years on average.³⁰ From the 15th day of the patients' cycles evening primrose oil or a not defined placebo was administered up to the beginning of the next cycle. Half of the patients started using evening primrose oil, the other half used a placebo. Allocation to groups was random. The length of the trial was 4 cycles. The symptoms were scored on a 3-point-scale before the trial, and during the trial on the last day of treatment. The sum of the scores formed the PMS-score, in which the three major symptoms counted twice (tender breasts, swelled breasts and a puffed feeling in the abdomen). The mean score for these three symptoms before treatment was 16.6, and after treatment 12.0 and 13.4 for the evening primrose oil-group and the placebo group respectively. As far as depressions are concerned the evening primrose oil-group scored slightly better than the placebo group. The scores for various other symptoms separately and the total score of all symptoms together did not show any differences. More patients in the evening primrose oil-group found that there was at least a slight improvement of their symptoms (62%), as compared with the placebo treatment (40%).

Brush, and Öckerman too, published other controlled studies on this disease as an abstract in which a clear reduction of the symptoms is mentioned.^{12,19}

No further controlled trials are known to us, and we think that the above mentioned trials form a still too unstable basis to propagate evening primrose oil as a therapy in PMS.

Related to PMS are complaints about pain and tension in the breasts. Pye et al. mentioned in a disorderly report the results of "randomised trials and open studies in 291 patients."³¹ The conclusion was that for several interventions good or useful results were achieved. For evening primrose oil and bromocriptine this was 45% and 47% respectively. In a therapy with danazol in patients with cyclic complaints of mastalgia this was the case for 70%. A well-executed trial would be interesting, as evening primrose oil does not show clear side-effects.

Other indications

McIllmurray et al. examined the effect of 6 capsules of evening primrose oil daily in patients with colorectal cancer in the Dukes C stage (no metastases, positive glands), who had been operated less than one month before this treatment.³² The period of administration was unlimited. Daily they also received 750mg of vitamin C, 150mg of pyridoxine and 30mg of zinc-sulphate. Soon after the beginning of the treatment four patients refused further co-operation and one patient stopped the treatment after 12 months. Of the other 49 patients 25 received evening primrose oil and 24 a placebo (randomly). Of these patients 10 and 12 respectively had died after a survival period of -median- 12 months for the two groups. This study was criticized for the fact that too small a dosage of GLA had been used, namely 270mg daily. In cases with a positive effect a dose was used that was 6 times larger.³³ McIllmurray et al. reacted to this by stating that long term swallowing of for example 36 capsules daily did not seem reasonable to them.

Jamal et al. carried out a study of the effect of evening primrose oil on the symptoms of diabetic neuropathy.³⁴ In this double-blind, randomized and placebo controlled trial in 22 patients (12 were treated with evening primrose oil, 8 capsules daily during 6 months,

and 10 were treated with a placebo), on which only a brief report appeared, a positive effect was found in all variables tested compared with the placebo. The only conclusion that one can draw from this study is, in our opinion, that it would be interesting to carry out more extensive trials.

Glen et al. examined the effect of 8 capsules of evening primrose oil daily during three weeks in alcohol dishabituation in which 62 patients received evening primrose oil and 59 controls received capsules containing paraffine.³⁵ Moreover, all the patients received intravenous vitamins during three days, and vitamin B complex supplements during the entire period of three weeks. They also received diazepam when they wanted, which was used as an end point too. Patients were allocated to groups via a "quota allocation system". A complete scale of assessments was performed with the result that there was a trend towards better outcomes for the evening primrose oil-group. Twenty-four patients had withdrawn, 20 of whom stopped on account of dismissal (13 in the active group and 7 in the placebo group) and 4 on account of "relapse or other psychiatric complication". This makes it difficult to interpret the results, although the evening primrose oil could be credited for a positive effect.

Wolkin et al. gave their patients with tardive dyskinesia 600mg of GLA or a placebo for six weeks on end.³⁶ This small trial with two groups of eight patients, which was concisely but orderly reported, did not show any differences between the two groups on the "Abnormal Involuntary Movements Scale" and on the "Brief Psychiatric Rating Scale". Evening primrose oil has also been tested as an antiobesity agent. Haslett et al. did not find any positive results.³⁷

Aman et al., in hyperactive children, could not find any positive changes on evening primrose oil.³⁸

Belch et al. examined the effect of 12 capsules of evening primrose oil daily on the length, the frequency and the severity of bouts in 11 patients and 10 controls, all of them with Raynaud's syndrome or with Raynaud's disease. The results were positive for the intervention group for 3 out of 4 end points used.^{39,40}

In general it can be stated that the number of studies undertaken is too small to arrive at definitive conclusions.

Discussion

The first thing that strikes one is the large number of disorders in which evening primrose oil was tested. There are several obvious explanations. Prostaglandins and related substances affect a scale of disorders, and a medicine that affects the formation of these substances could, theoretically, be useful in many diseases. The developments in this specific area have gone very rapidly during the last decade, and warrant further clinical studies.

For most disorders that have been investigated there is no satisfactory "regular" therapy, and the administration of evening primrose oil does not seem to cause negative side-effects. A medicine that would have so many positive effects is of course a gold-mine for the industry, certainly because it completely meets the present boom in the use of "natural medicines." The limited presence of GLA in foods is almost a guarantee for commercial success. In that case studies will have to be of a good quality in order to obtain valid results. As said before, the problems are related to this matter. Besides factors such as sufficiently large numbers of patients and an adequate follow-up, problems that are harder to solve also play a part. In order to guarantee double blinding a well matched placebo must be used. In most trials we reviewed capsules containing paraffine were used. We have seen some of these capsules, and they matched the evening primrose capsules by colour and form. Their taste was different, but the capsules must be swallowed unbro-

ken. Unfortunately the paraffine capsules that were shown to us did not have the logo that was on the real capsules.

The composition of the capsules is a problem as well. They contain linoleic acid, GLA, and 10 IE of vitamin E. In doses of for example 12 capsules daily, the level of vitamin E is a megadose level (10 times the recommended daily allowance in this example). One can wonder in such a case which substance is the substance investigated, especially since in most of the above-mentioned trials it was not reported whether vitamin E had been added to the paraffine capsules. We suggest as a placebo capsules containing linoleic acid and vitamin E so that the effect of GLA (which is credited for the effect) can be assessed.

Quite a lot of trials are so-called cross-over trials. One of the starting points in this kind of trials is that the groups must be comparable prognostically at the time of the cross-over. This is only possible in certain chronic diseases with a predictable course, in which after a so-called wash-out period the effects of the first intervention have disappeared. Only in the trial by Aman et al.³⁸ a wash-out period of a week is reported. The reports of the other trials do not mention a wash-out period. This would imply that the effects of evening primrose oil would disappear immediately after stopping its administration. When a sustained effect is supposed, there should be a wash-out period, otherwise the substance that causes the effect is hard to assess (carry-over effect), and the results of these trials would be invalid.⁴¹ One could look at the results of the first half of the trial but, unfortunately, these data are reported seldomly.

The next problem is the assessment of the compliance. Mostly, no data about the compliance are reported, but it seems reasonable to assume that it has not been optimal, see for example the trial of Bamford et al. With an active agent bad compliance will cause a decrease in the effect so that it becomes more difficult to demonstrate it. If positive effects are found, these could be larger in the ideal situation.

The optimal dosage is often not known. Trials that tried to establish a dose-response-effect show varying results.^{28,29} The dosage administered varies from 2 to 16 capsules daily in the trials presented here. In the leaflet in medicines 4 to 6 capsules daily are recommended. In practice it will be difficult to motivate patients to take many capsules daily. The optimal length of the intervention is also mostly unknown. The recurrence of symptoms after stopping the use of evening primrose oil does not seem hypothetical. The follow-up after stopping the intervention is a matter to which no attention is paid in many trials, and the above-mentioned trials are no exception at this point.

When we look at the results, weighing all items discussed, we conclude that for disorders like Raynaud's phenomenon, premenstrual syndrome, and Sjögren's syndrome further studies are warranted, preferably in comparison with current therapies. If the GLA really causes the effects, a cheaper and more efficient way of administering it than through evening primrose oil would be desirable. We are thinking, for example, of GLA in a pure form (possibly manufactured synthetically), with which much larger dosages can be given.

Conclusions

Many studies have been undertaken to examine the efficacy of evening primrose oil. The trials that had clinical phenomena as end points generally showed quite a lot of methodological defects. At several indications further studies to examine the efficacy of evening primrose oil or GLA seem warranted, particularly because here disorders are concerned for which often no good therapy exists (yet). As long as the efficacy of evening primrose oil has not been proven, we do not recommend to use it in view of the quantity and the price of the capsules that should be taken.

References

1. Hornstra G. Nieuwe inzichten met betrekking tot essentiële vetzuren. *Voeding* 1987;48:326-30.
2. Horrobin DF. A new concept of lifestyle-related cardiovascular disease: the importance of interactions between cholesterol, essential fatty acids, prostaglandin E1 and tromboxane A2. *Med Hypotheses* 1980;6:785-800.
3. Horrobin DF, Huang Y-S. The role of linoleic acid and its metabolites in the lowering of plasma cholesterol and the prevention of cardiovascular disease. *Int J Cardiol* 1987;17:241-55.
4. Christie SBM, Conway N, Pearson HES. Observations on the performance of a standard exercise test by claudicants taking γ -linolenic acid. *J Atheroscler Res* 1968;8:83-90.
5. Lovell CR, Burton JL, Horrobin DF. Treatment of atopic eczema with evening primrose oil. *Lancet* 1981;1:278.
6. Pashby NL, Mansel RE, Hughes LE, Hanslip J, Preece PE. A clinical trial of evening primrose oil in mastalgia. *Br J Surg* 1981;68:801.
7. Vapaatalo H, Jäntti J, Isomäki H, Nikkari T, Seppälä E. Treatment of rheumatoid arthritis with linolenic acid (LA). *Scand J Rheumatol* 1983;49:16.
8. Chalmers RJG, Shuster S. Evening primrose seed oil in ichthyosis vulgaris. *Lancet* 1983;1:236-7.
9. Holman CP, Bell AFJ. A trial of evening primrose oil in the treatment of chronic schizophrenia. *Journal of Orthomolecular Psychiatry* 1983;12:302-4.
10. London RS, Sundaram GS, Murphy L, Goldstein PJ. Evaluation and treatment of breast symptoms in patients with the premenstrual syndrome. *J Reprod Med* 1983;28:503-8.
11. Manthorpe R, Hagen Petersen S, Prause JU. Primary Sjögren's syndrome treated with Efamol/Efavit. A double-blind cross-over investigation. *Rheumatol Int* 1984;4:165-7.
12. Brush MG, Watson SJ, Perry M, et al. Gamma-linolenic acid in the treatment of premenstrual syndrome. 2nd International congress on essential fatty acids, prostaglandins and leukotrienes. London, 24-27 March 1985: abstract 15.
13. Macdonald KJS, Green C, Raffle EJ, Kenicer KJA. Topical evening primrose seed oil and atopic eczema. *Scott Med J* 1985;30:267.
14. Simpson LO, Shand BI, Olds RJ. Dietary supplementation with Efamol and multiple sclerosis. *NZ Med J* 1985;98:1053.
15. Ylikorkala O, Puolakka J, Mäkääinen L, Viinikka L. Prostaglandins and premenstrual syndrome. *Prog Lipid Res* 1986;25:433-5.
16. Skogh M. Atopic eczema unresponsive to evening primrose oil (linoleic and γ -linolenic acids). *J Am Acad Dermatol* 1986;15:114-5.
17. Ansell D, Belch JFF, Curan L, McLaren M, O'Dowd A, Forbes CD. The effects of Efamol and Efamol Marine on patients with rheumatoid arthritis: a double blind, placebo controlled study. 6th International congress on prostaglandins. Florence, June 1986:24.
18. Biagi PL, Bordoni A, Masi M, et al. Use of evening primrose in children with atopic eczema [abstract]. IX International symposium on drugs affecting lipid metabolism. Florence, 22-25 October 1986.
19. Öckerman PA, Bachrack I, Glans S, Rassner S. Evening primrose oil as a treatment of the premenstrual syndrome. *Recent Advances in Clinical Nutrition* 1986;2:404-5.
20. Schalin-Karrila M, Mattila L, Jansen CT, Uotila P. Evening primrose oil in the treatment of atopic eczema: effect on clinical status, plasma phospholipid fatty acids and circulating blood prostaglandins. *Br J Dermatol* 1987;117:11-9.
21. Bates D, Fawcett PRW, Shaw DA, Weightman D. Trial of polyunsaturated fatty acids in non-relapsing multiple sclerosis. *Br Med J* 1977;2:932-3.
22. Bates D, Fawcett PRW, Shaw DA, Weightman D. Polyunsaturated fatty acids in treatment of acute remitting multiple sclerosis. *Br Med J* 1978;2:1390-1.
23. Bates D, Fawcett PRW, Shaw DA, Weightman D. Polyunsaturated fatty acids in multiple sclerosis. *Br Med J* 1979;1:683.
24. Dworkin RH, Bates D, Millar JHD, Paty DW. Linoleic acid and multiple sclerosis: A reanalysis of three double-blind trials. *Neurology* 1984;34:1441-5.
25. Belch JFF, Ansell D, Madhok R, O'Dowd A, Sturrock RD. Effects of altering dietary essential fatty acids on requirements for non-steroidal anti-inflammatory drugs in patients with rheumatoid arthritis: a double blind placebo controlled study. *Ann Rheum Dis* 1988;47:96-104.
26. Krakauer K, Torrey SB, Zurier RB. Prostaglandin E1 treatment of NZB/W mice. *Clin Immunol Immunopathol* 1978;11:256-64.

27. Oxholm P, Manthorpe R, Prause JU, Horrobin D. Patients with primary Sjögren's syndrome treated for two months with evening primrose oil. *Scand J Rheumatol* 1986;15:103-8.
28. Wright S, Burton JL. Oral evening-primrose-seed oil improves atopic eczema. *Lancet* 1982;2:1120-2.
29. Bamford JT, Gibson RW, Renier CM. Atopic eczema unresponsive to evening primrose oil (linoleic and γ -linolenic acids). *J Am Acad Dermatol* 1985;13:959-65.
30. Puolakka J, Mäkääinen L, Viinikka L, Ylikorkala O. Biochemical and clinical effects of treating the premenstrual syndrome with prostaglandin synthesis precursors. *J Reprod Med* 1985;30:149-153.
31. Pye JK, Mansel RE, Hughes LE. Clinical experience of drug treatments for mastalgia. *Lancet* 1985;2:773-7.
32. McIlmurray MB, Turkie W. Controlled trial of γ -linolenic acid in Dukes's C colorectal cancer. *Br Med J* 1987;294:1260. Correction *Br Med J* 1987;295:475.
33. Van der Merwe CF. Controlled trial of γ -linolenic acid in Dukes's C colorectal cancer. *Br Med J* 1987;295:852.
34. Jamal GA, Carmichael H, Weir AI. Gamma-linolenic acid in diabetic neuropathy. *Lancet* 1986;1:1098.
35. Glen AIM, Glen EMT, MacDonald FK, MacDonell LEF, MacKenzie JR, Montgomery DJ. Essential fatty acids in the treatment of the alcohol dependence syndrome. In: Birch GG, Lindley MG, eds. *Alcoholic Beverages*. London: Elsevier Science Publishers, 1985:203-21.
36. Wolkin A, Jordan B, Peselow E, Rubinstein M, Rotrosen J. Essential fatty acid supplementation in tardive dyskinesia. *Am J Psychiatry* 1986;143:912-4.
37. Haslett C, Douglas JG, Chalmers SR, Weighhill A, Munro JF. A double-blind evaluation of evening primrose oil as an antiobesity agent. *Int J Obesity* 1983;7:549-53.
38. Aman MG, Mitchell EA, Turbott SH. The effects of essential fatty acid supplementation by Efamol in hyperactive children. *J Abnorm Child Psychol* 1987;15:75-90.
39. Belch JFF, Shaw B, O'Dowd A, et al. Evening primrose oil (Efamol) in the treatment of Raynaud's phenomenon: a double blind study. *Thromb Haemost* 1985;54:490-4.
40. Belch JFF, Shaw B, O'Dowd A, et al. Evening primrose oil (Efamol) as a treatment for cold-induced vasospasm (Raynauds phenomenon). *Prog Lipid Res* 1986;25:335-40.
41. Hills M, Armitage P. The two-period cross-over clinical trial. *Br J Clin Pharmacol* 1979;8:7-20.

Garlic, onions and cardiovascular risk factors.

A review of the evidence from human experiments with emphasis on commercially available preparations.

J. KLEIJNEN, P. KNIPSCHILD & G. TER RIET

Department of Epidemiology/Health Care Research, University of Limburg, Postbox 616, 6200 MD Maastricht, The Netherlands

- 1 Claims for beneficial effects on cholesterol levels, fibrinolytic activity, and platelet aggregation are attributed both to fresh garlic and onions (or their extracts) and to commercially available preparations.
- 2 Regarding fresh garlic, the claims have been confirmed, but so far only at very high dosages.
- 3 For onions and commercially available supplements contradictory results have been reported.
- 4 All published trials showed severe methodological shortcomings. Some trials were not randomized and/or not blinded whilst this was possible, and in only one of every three studies more than 25 patients participated in each treatment group. In no trial was prognostic comparability of the treatment and the control groups ascertained. At the moment there is inadequate scientific justification for garlic supplementation.

Keywords garlic onions commercial preparations cardiovascular disease

Introduction

People who take garlic (*Allium sativum*) supplements hope to improve their well-being and to reduce the risk of various diseases. Numerous myths about its medicinal properties have existed for centuries in various cultures. According to some investigators the lower incidence of cardiovascular disease in southern countries may partly be due to consumption of large quantities of garlic (Slater, 1979). Some more specific claims have been made on the beneficial influence of garlic on several risk factors for cardiovascular disease. Garlic is said to lower serum cholesterol, enhance blood fibrinolytic activity, and to inhibit platelet aggregation. Additional claims have been made on its hypoglycaemic effects, antibiotic actions and anticarcinogenic properties. Similar claims have

appeared for onions (*Allium cepa*). Thousands of publications have appeared in the last 20 years dealing with research on garlic and onions.

We found 18 controlled trials about the effects on presumed cardiovascular risk indicators in humans. In this article emphasis is put on the results of trials on commercial garlic preparations. Comparing the effects of these preparations to fresh garlic is interesting, especially for those preparations which are (or claim to be) odourless. The typical odour is an unavoidable side effect and substantial amounts of fresh garlic seem to be needed to obtain measurable effects. This review deals critically with methodological problems of these trials. Besides an assessment of the evidence, suggestions for further research are given.

Biochemical aspects

Alliin ((+)-S-allyl-L-cysteine sulphoxide) is converted by action of the enzyme allinase into allicin (diallyl thiosulphate). This happens when the garlic bulb is crushed. Allicin can be converted (by heat) into diallyldisulphide which in its turn is converted into various sulphide containing substances which cause the typical smell of garlic (di-, tri-, polysulphides). Allicin and diallyldisulphide combine to 4,5,9-trithiadodeca-1,6,11-triene 9-oxide which is called ajoene. In onions the same and related substances are found e.g. cycloalliin (5-methyl-1,4-thiazan-3-carboxylic acid 1-oxide), which is odourless (Augusti & Benaim, 1975). Whitaker (1976) gives a list of compounds of intact onion, of crushed onion and crushed garlic, and a list of compounds identified in steam-distilled onion oil. Additional compounds have been reported by several authors (Ariga *et al.*, 1981; Block & Ahmad, 1984; Bayer *et al.*, 1988). Block & Ahmad (1984) suggest a scheme by which the decomposition of allicin takes place.

Certain actions (on blood lipids, coagulation factors, fibrinolytic activity, blood sugars and antibiotic actions) of onion and garlic are ascribed to specific components listed in Table 1, which is not intended to be complete.

Methods

Experiments were found by a computer search (MEDLINE CD-ROM 1983–1989). Further experiments were found by checking references extensively and by personal communication with authors of experiments. Trials were only reviewed for this article if a control group was included. Only experiments in humans are discussed.

For any claim of therapeutic effects of a specified substance, a randomized, double-blind, placebo-controlled trial is the method of preference to assess specific effects attributable to the experimental intervention. Randomization (and restriction of the test group and/or prestratification if appropriate) takes care of known and unknown confounders which otherwise could influence the prognostic comparability of the test groups. In trials on the effects of garlic or onions, randomization should be no problem. The substance used in the trial, its preparation, its dosage and the way in which it is taken by the patients should be described adequately. Blinding is a major problem in garlic trials because of its characteristic smell. If fresh, cooked or fried

garlic is used, blinding the patient is not possible. Only trials of a very short duration can be done with these substances because of problems in controlling the external circumstances for both groups. In several trials garlic was given in this form for weeks or months. Controlling of the patient's diets and other habits for such a long period is an illusion, and differences in uncontrolled factors may account for the results. If capsules or tablets are used, placebo capsules or tablets, indistinguishable from the real ones, should be given. People should be asked if they know in what treatment group they are, to check the blinding. Also compliance must be checked.

Measurements of the effects, especially blood tests of cholesterol levels, fibrinolytic activity and tests of platelet aggregation, should be performed with accepted methods, and physiological variations should be taken into account. For instance, blood fibrinolytic activity may vary with different blood fibrinogen concentrations and other factors.

Statistical analysis in controlled trials should be based on 'between groups comparisons' and differences should be presented with confidence intervals. In most articles presented in this review the reported results and significance levels were based on 'within group comparisons'. Differences between groups with confidence intervals cannot be given because of lack of data in the articles themselves.

Results

The results presented in the tables show the percentage of change from baseline levels for treatment and control groups.

Fresh garlic

Five out of six trials in which fresh garlic or freshly made extracts were used, and in which cholesterol was measured, show a lowering of serum cholesterol. Furthermore, it was shown that garlic increases fibrinolytic activity and inhibits platelet aggregation in all studies reporting these measurements. Unfortunately, the dosages needed to obtain these effects are relatively high. The dosage which was mostly used is approximately (the equivalent of) 0.25 to 1 g of fresh garlic kg^{-1} bodyweight (7 to 28 cloves each day!).

Fresh onions

An increase of fibrinolytic activity was found in all three trials in which onions were used, while

Table 1 Active principles of garlic and onions**Anti-platelet:***Garlic:*

- alliin (Hanley & Fenwick, 1985)
- allicin (Block & Ahmad, 1984)
- allyl-1.5-hexadienyl-trisulphide (Block & Ahmad, 1984)
- allyl methyl trisulphide (Block & Ahmad, 1984)
- S-allyl 2-propene thiosulphinat (Block & Ahmad, 1984)
- ajoene (Apitz-Castro *et al.*, 1984)
- diallyl disulphide (Ariga *et al.*, 1981)
- diallyl trisulphide (Ariga *et al.*, 1981)
- 1.5-hexadienyl-trisulphide (Apitz-Castro *et al.*, 1983)
- methyl allyl trisulphide (Ariga *et al.*, 1981)
- 2-vinyl-1.3-dithiene (Apitz-Castro *et al.*, 1983)
- 3-vinyl-1.2-dithiene (Block, 1985)

Onions

- adenoside (Kawakishi & Morimitsu, 1988)
- alliin (Liakopoulou-Kyriakides, 1985)
- 1-(methyl sulphinyl)-propyl methyl disulphide (Kawakishi & Morimitsu, 1988)
- 9,10,13-trihydroxy-11-octadecenoic acid (Üstünes *et al.*, 1985)
- 9,12,13-trihydroxy-10-octadecenoic acid (Üstünes *et al.*, 1985)
- trans-5-ethyl-4,6,7-trithia-2-decene-4-oxide (Bayer *et al.*, 1988)
- trans,trans (and trans,cis) 5-ethyl-4,6,7-trithia-2,8-decadiene-4-oxide (Bayer *et al.*, 1988)

Antibiotic*Garlic:*

- allicin (Hanley & Fenwick, 1985)
- ajoene (Yoshida *et al.*, 1987)
- diallyl disulphide (Amonkar & Reeves, 1971)
- diallyl trisulphide (Amonkar & Reeves, 1971)

Fibrinolysis:*Garlic:*

- methane-thiol-3,4-dimethylthiophene (Augusti & Benaim, 1975)
- methyl cysteine sulfoxide (Augusti & Benaim, 1975)
- propyl allyl disulphide (Augusti & Benaim, 1975)
- propyl cysteine sulfoxide (Augusti & Benaim, 1975)

Onions:

- cycloalliin (August & Benaim, 1975)

Blood sugar, insulin:*Garlic:*

- allicin (Hanley & Fenwick, 1985)
- diallyl disulphide (Hanley & Fenwick, 1985)

Onions:

- allyl propyl disulphide (Augusti & Benaim, 1975)

Blood lipids, cholesterol:*Garlic:*

- alliin (Kamanna & Chandrasekhara, 1984)
- allicin (Kamanna & Chandrasekhara, 1984)
- allyl propyl disulphide (Bordia, 1975)
- diallyl disulphide (Bordia *et al.*, 1975)
- S-methyl-L-cysteine sulfoxide (Kamanna & Chandrasekhara, 1984)

Table 2 Results of trials in which fresh garlic or freshly made extracts were used, as percentage of change compared with baseline

Author, year, randomization	Number, fat load	Preparation, dosage	Duration	Cholesterol	Fibrinolytic activity	Coagulation platelets
Bordia <i>et al.</i> (1975), R	10 per group, butter	garlic (d,e), equal to 50 g	once	(+) -6; -7/7	(+) 15; 15/-49	(+) 23; 22/-13 coagulation time
Bordia <i>et al.</i> (1977)	10/10	garlic (d), equal to 1g kg ⁻¹	20 days	NM	(+) 96/24	NM
Bhushan <i>et al.</i> (1979) R	15/10	garlic (a), 10 g	2 months	(+) -15/0	NM	NM
Sucur (1980) R	200, crossover	garlic (a,f), 15 g; 30 g	25 days	(+) -10; -10/9	NM	NM
Bordia (1981) R	33/29	garlic (d), 0.25 mg kg ⁻¹ (equal to 30 g)	10 months	(+) -18/0	NM	NM
Chutani & Bordia (1981) R	10; 10/10	garlic (a,b), 0.5 g kg ⁻¹	1 month	NM	(+) 85; 72/3	NM
Arora & Arora (1981)	30, crossover, butter	garlic (d), equal to 50 g; clofibrate, 500 mg	once	(-) 0; 1/-1	(+) 3; 71/-31	(+) 0; -2/-18, coagulation time
Bordia <i>et al.</i> (1982)	20/20, butter	garlic (d), 0.25 mg kg ⁻¹	3 weeks	(+) 2/10	(+) 36/-23	(+) -16/31, platelet adhesive index

R = randomized
 (+) = beneficial effect
 (-) = no effect
 NM = not measured
 / = separation of treatment group(s) from control group(s)
 ; = separation of different treatment groups
 a = fresh garlic
 b = fried garlic
 c = steam distilled extract of garlic
 d = ether extract of garlic
 e = juice of garlic
 f = water-ethanol extract

Table 3 Results of trials in which fresh onions or freshly made extracts were used, as percentage of change compared with baseline

Author, year, randomization	Number, fat load	Preparation, dosage	Duration	Cholesterol	Fibrinolytic activity	Coagulation platelets
Menon <i>et al.</i> (1968)	8; 14, crossover, breakfast	onions (a,b), 60 g	once	(-) a,b pooled, 3/-2	(+) 50; 47/-14; -25	(-) a,b pooled, 0/3; recalcified clotting time
Bordia <i>et al.</i> (1975), R	10; 10/10, butter	onions (c,d), equal to 50 g	once	(+) -3; -3/7	(+) 16; 16/-49	(+) 11; 10/-13, coagulation time
Arora & Arora (1981)	30, crossover, butter	onions (c), equal to 50 g; clofibrate 500 mg	once	(-) -2; 1/-1	(+) 0; 71/-31	(-) -14; -2/-18, coagulation time

R = randomized

(+) = beneficial effect

(-) = no effect

NM = not measured

/ = separation of treatment group(s) from control group(s)

; = separation of different treatment groups

a = boiled onions

b = fried onions

c = ether extract of onions

d = juice of onions

measurements of cholesterol levels and platelet aggregation showed contradictory results.

Sharma *et al.* (1977) found in a small trial that an aqueous extract of onions had hypoglycaemic effects. Other parameters have to our knowledge not been tested in humans with control groups. For further details of the studies in which fresh garlic and onions were used and for the results of uncontrolled human studies and of interesting animal studies, we refer to other reviews (Lau *et al.*, 1983; Fenwick & Hanley, 1986; Ernst, 1987; Kendler, 1987). A serious side effect – the odour – invariably occurs at effective dosages of garlic, which will be unacceptable for most people.

Commercially available preparations

What about the commercially available (de-odourized) products? In studies with rats and pullets deodourized garlic was shown to retain active components (Qureshi *et al.*, 1983; Lau *et al.*, 1987). There were only seven controlled studies in humans.

Augusti & Benaim (1975) compared the effects of a commercially produced steam distilled extract of onions with self-prepared ether extracts and to purified cyclo-alliin on blood fibrinolysis. The steam-distilled oil yielded 0.28 g kg^{-1} of raw onion. The dosage was 125 mg per subject (healthy volunteers) given once. The ether-extractable oil yielded 1 g kg^{-1} of raw onion. Cyclo-alliin (yield 0.2 g kg^{-1}) was given in a dose of 0.125 g. A number of sulphur-containing components was present in both extracts, as was shown by gas chromatography. Measurements were performed before and 2 h after the products had been taken. This study, which was not randomized and not blinded and in which six patients in each group participated, showed that all three substances increased the fibrinolytic activity compared with controls.

Agarwal *et al.* (1977) assessed the effect of 0.25 g synthesized cyclo-alliin (equal to 1.2 kg of extracted onions) or placebo (lactose) on fibrinolytic activity and platelet aggregating activity in 18 male volunteers in a randomized double-blind cross-over trial. Measurements were performed before and 1.5 h after the medication was taken. They found an increased fibrinolytic activity and no effect on platelet aggregability.

Lutomski (1984) tested in a double-blind randomized study the effect of six garlic pills a day (Ilja Rogoff garlic pills with rutin, one pill contains 50 mg 'bulb. allii sat sicc.' and other substances) against undefined placebos in 102 patients aged 45–60 years complaining of de-

creased operating capacity, joy of life and physical well-being. After 12 weeks full data were available of 82 patients, 44 on garlic and 38 on placebo. Lipids, fibrinolytic activity, psychologic tests, pulse, and glucose showed no differences. Positive effects were measured for subjective symptoms (headache, sleep, dizziness, digestion, joy of life, physical well-being, operating capacity) and blood pressure. For all subjective symptoms 52% of patients using garlic showed improvement, compared to 35% of the patients using placebo. Data on blood pressure were only given for patients with initial high pressures (22 using garlic, 14 on placebo). Improvement was found in 17 patients on garlic and eight patients on placebo. It was not mentioned whether compliance or blinding was checked.

Ernst *et al.* (1985) treated 20 hypercholesterolaemic patients in an open study with a hypocaloric diet. Ten patients also received 21 g garlic powder (Kwai Tabs) equivalent to 64 g fresh garlic a day for 4 weeks. They found lower cholesterol and triglyceride levels in the patients treated with garlic.

Luley *et al.* (1986) assessed the effects of a commercial dried garlic preparation in patients with primary hyperlipoproteinemia types IIa, IIb and IV in two randomized double-blind cross-over studies. Garlic and placebo treatment were given for 6 weeks each. Dried garlic was administered in a dosage of 600 mg a day to 34 patients in the first study, and in a dosage of 1350 mg a day to 51 patients in the second study. The higher dose corresponds to 5 g of fresh garlic. Compliance was estimated 86% by tablet counts. Neither dosage showed any effect on many lipid parameters and coagulation parameters. The use of the relatively low dosages compared to those administered in other studies, and a possible loss of the active principle due to the drying process might explain the negative results according to Luley *et al.* We would like to add that a carry-over effect might explain the negative results since no wash-out period was mentioned. Some of the patients in the second study reported complaints of a bad smell.

Lau *et al.* (1987) found in a randomized study that a liquid garlic extract (Kyolic) was effective in lowering serum cholesterol and triglycerides in hyperlipidaemic patients. Sixteen patients (initial cholesterol $220\text{--}440 \text{ mg dl}^{-1}$, conversion factor to mmol l^{-1} is 0.026) took four 1 ml capsules containing 250 mg ml^{-1} dry weight of the active garlic components and 16 patients took an indistinguishable placebo for 6 months. Fourteen patients were lacto-ovo vegetarians, 10 were occasional meat eaters and eight were

regular meat eaters. It was not stated whether the patients and investigators were blinded. Five patients withdrew, one taking garlic and four taking placebo, because of minor discomforts. After 6 months mean serum cholesterol had changed from 306 to 262 mg dl⁻¹ and from 302 to 292 mg dl⁻¹ in the garlic and placebo group respectively. Data on triglycerides, HDL, LDL etc. were only shown for the patients taking garlic. In normolipidaemic patients (fourteen subjects receiving garlic or placebo) 'data . . . are not shown since no significant changes were noted'.

Sitprija *et al.* (1987) studied the effect of 350 mg of garlic (spray dried method) or an undefined placebo twice daily in a double-blind randomized study in forty non-insulin dependent diabetic patients. Blood glucose and serum immunoreactive insulin responses to an oral glucose load were measured before and 1 month after treatment. Blood total cholesterol, high density lipoprotein and triglyceride levels were

also measured. Thirty-three patients completed the study. No differences, and no side effects of garlic were observed. It was not mentioned whether compliance or blinding was checked.

Discussion and conclusion

Commercial garlic preparations are often manufactured with heat application in at least some stage of the process. Besides heat, drying and pulvering, the essential oil is often extracted by steam distillation using ether or methanol. At least one company (Wakunaga, Kyolic) uses a so-called 'cold aging process' which means that the garlic bulbs are stored for some 20 months. In this manner components that cause the bad smell are converted or vapourized. Whether active substances are left after the processing and storage is of course the crucial question. The exact composition of commercial garlic preparations is mostly unknown and surely there are

Table 4 Results of trials in which commercially available preparations were used, as percentage of change compared with baseline

Author, year, randomization	Number, fat load	Preparation, dosage	Duration	Cholesterol	Fibrinolytic activity	Coagulation platelets
<i>Onions</i>						
Augusti & Benaim (1975)	6/6	steam dist.	once	NM	(+) 87/7	NM
		ether dist.; 2 preparations	once	NM	(+) 68;50/13;15	NM
		cycloalliin	once	NM	(+) 42/13	NM
Agarwal <i>et al.</i> (1977), R	18 crossover	cycloalliin	once	NM	(+) 77/7	(-) no difference
<i>Garlic</i>						
Lutomski (1984), R	44/38	Ilja Rogoff	12 weeks	(-) no diff.	(-) no diff.	NM
Ernst <i>et al.</i> (1985)	10/10	Kwai tabs	4 weeks	(+) -22/-12	NM	NM
Luley <i>et al.</i> (1986), R	34;51 crossover	dried garlic	6 weeks	(-) no diff.	NM	(-) no diff.
Lau <i>et al.</i> 1987, R	16/16	Kyolic	6 months	(+) -14/-3	NM	NM
Sitprija <i>et al.</i> (1987), R	20/20	spray dried garlic	1 month	(-) 5/1	NM	NM

R = randomized

(+) = beneficial effect

(-) = no effect

NM = not measured

/ = separation of treatment group(s) from control group(s)

; = separation of different treatment groups

many preparations on the market in which active compounds like ajoene and the cyclic anti-thrombotic compounds cannot be isolated (Block, 1985). Most reports mention the laboratory methods used, but the contents of the preparations of garlic or onions are not described (except in the studies in which cyclo-alliin was used). One should bear in mind that the dose of the different components in garlic may vary to a great extent in different regions and years, which might influence the composition of garlic preparations.

The results of the experiments with fresh garlic are consistent: garlic causes an increase of fibrinolytic activity, it inhibits platelet aggregation and it also lowers cholesterol levels. The dosages needed are for most western people unacceptably high: at least 7 cloves of garlic each day. We did not find experiments in which low dosages (for instance one clove a day) were used. In several (cross-over) experiments four or five different preparations were given once in the same week. At each following day, initial levels of cholesterol and fibrinolytic activity had returned to baseline levels. Therefore, it seems that the effects only last for several hours after intake. Bordia (1981) and Lau *et al.* (1987) in longer term studies reported an initial raise of cholesterol levels (Bordia used an ether extract of garlic and Lau *et al.* used Kyolic) and a decrease below baseline levels after several months of garlic intake. Fresh garlic may cause dermatitis, gastrointestinal problems and allergic reactions.

Fresh onions also appear to cause an increase of fibrinolytic activity but regarding the other parameters contradictory results have been reported.

When commercially available preparations were tested, contradictory results have been reported too. Besides a bad smell, no side effects of commercial preparations were reported.

It is easy to criticize a study on methodological shortcomings. In practice, however, it is often not possible to conduct an impeccable trial. When fresh garlic is used, blinding of the patient is not possible. This may result in differences in external circumstances. The risk of measurement bias is limited when only laboratory parameters are assessed. The results of these unblinded trials may be biased, but they show a remarkable consistency.

Only eleven out of eighteen trials were randomized, five in which fresh garlic was used, one in which onions were used and five in which commercial preparations were used, but not one

described how the randomization was executed. From the randomized studies using commercial preparations four out of five were blinded. The number of people in the studies was limited. In only six studies more than 25 patients participated in each treatment group. In no article was it mentioned whether the placebo was tested for indistinguishability and whether blinding was checked when appropriate. Compliance was checked in two studies. Proper randomization, sufficient numbers of patients, checking of the placebo, compliance and blinding are factors that must be controlled whenever possible. It can be concluded that all trials of the effects of garlic and onions on factors associated with cardiovascular disease show several major methodological shortcomings.

Several of the active components such as allicin, ajoene and cycloalliin (found in onions) can be synthesized. Many components of garlic and onion have been identified, and supplementation with isolated components of garlic may give no side effects, while causing beneficial effects. Unfortunately, the components that cause the bad smell may also be the active ones i.e. sulphides, with one or two exceptions like cycloalliin.

If a substance is effective in lowering serum cholesterol and if it inhibits platelet aggregation it does not necessarily follow that it is good for general health. It is doubtful whether healthy people need a change of their fibrinolytic activity, or inhibition of their platelet aggregation. Even lowering cholesterol levels is no guarantee for beneficial effects (Committee of Principal Investigators, 1978). Trials with endpoints that are more relevant than mere laboratory measurements are necessary before claiming that garlic improves health. For instance, morbidity and mortality rates could be assessed. These trials are, of course, expensive and time consuming and there should be good reasons to start them. It is not probable that these trials will be done in the near future. In that case we suggest that the proper dose and duration of intake are assessed, and that the preparations of garlic and onions are standardized. Their effects should be compared to those of other drugs with similar action, before large scale trials are being considered. Side effects should also be taken into account.

For this moment there is inadequate scientific justification for garlic supplementation.

This article was supported by a grant of the Dutch Ministry of Welfare, Public Health and Cultural Affairs.

References

- Agarwal, R. K., Dewar, H. A., Newell, D. J. & Das, B. (1977). Controlled trial of the effect of cycloallin on the fibrinolytic activity of venous blood. *Atherosclerosis*, **27**, 347-51.
- Amonkar, S. V. & Reeves, E. L. (1971). Mosquito control with active principle of garlic. *J. Economic Entomology*, **63**, 1172-75.
- Apitz-Castro, R., Cabrera, S., Cruz, M. R., Ledezma, E. & Jain, M. K. (1983). Effects of garlic extract and of three pure components isolated from it on human platelet aggregation, arachidonate metabolism, release reaction and platelet ultrastructure. *Thrombosis Res.*, **32**, 155-69.
- Apitz-Castro, R., Escalante, J., Vargas, R. & Jain, M. K. (1986). Ajoene, the antiplatelet principle of garlic, synergistically potentiates the antiaggregatory action of prostacyclin, forskolin, indomethacin and dipyridamole on human platelets. *Thrombosis Res.*, **42**, 303-11.
- Ariga, T., Oshiba, S. & Tamada, T. (1981). Platelet aggregation inhibitor in garlic. *Lancet*, **i**, 150-51.
- Arora, R. C. & Arora, S. (1981). Comparative effect of clofibrate, garlic and onion on alimentary hyperlipemia. *Atherosclerosis*, **39**, 447-52.
- Augusti, K. T. & Benaim, M. E. (1975). Effect of essential oil of onion (allyl propyl disulphide) on blood glucose, free fatty acid and insulin levels of normal subjects. *Clin. Chim. Acta*, **60**, 121-23.
- Bayer, T., Wagner, H., Wray, V. & Dorsch, W. (1988). Inhibitors of cyclo-oxygenase and lipooxygenase in onions. *Lancet*, **ii**, 906.
- Bhushan, S., Sharma, S. P., Singh, S. P., Agrawal, S., Indrayan, A. & Seth, P. (1979). Effect of garlic on normal blood cholesterol level. *Ind. J. Physiol. Pharmacol.*, **23**, 211-14.
- Block, E. (1985). The chemistry of garlic and onions. *Scientific Am.*, **252**, 94-114, 119.
- Block, E. & Ahmad, S. (1984). (E,2) Ajoene: a potent antithrombotic agent from garlic. *J. Am. chem. Soc.*, **106**, 8295-8296.
- Bordia, A. (1981). Effect of garlic on blood lipids in patients with heart disease. *Am. J. clin. Nutr.*, **34**, 2100-03.
- Bordia, A., Bansal, H. C., Arora, S. K. & Singh, S. V. (1975). Effect of the essential oils of garlic and onion on alimentary hyperlipemia. *Atherosclerosis*, **21**, 15-19.
- Bordia, A., Joshi, H. K., Sanadhya, Y. K. & Bhu, N. (1977). Effect of essential oil of garlic on serum fibrinolytic activity in patients with coronary artery disease. *Atherosclerosis*, **28**, 155-59.
- Bordia, A., Sharma, K. D., Parmar, Y. K. & Verma, S. K. (1982). Protective effect of garlic oil on the changes produced by 3 weeks of fatty diet on serum cholesterol, serum triglycerides, fibrinolytic activity and platelet adhesiveness in man. *Indian Heart J.*, **34**, 86-88.
- Chutani, S. K., Bordia, A. (1981). The effect of fried versus raw garlic on fibrinolytic activity in man. *Atherosclerosis*, **38**, 417-21.
- Committee of Principal Investigators. (1978). A co-operative trial in the primary prevention of ischaemic heart disease using clofibrate. *Br. Heart J.*, **40**, 1069-1118.
- Ernst, E., Weihmayr, T. & Matrai, A. (1985). Garlic and blood lipids. *Br. med. J.*, **291**, 139.
- Ernst, E. (1987). Cardiovascular effects of garlic (*Allium sativum*): a review. *Pharmatherapeutica*, **5**, 83-89.
- Fenwick, G. R. & Hanley, A. B. (1986). The genus *Allium*-P3. *CRC. Crit. Rev. Food Sci. Nutr.*, **23**, 1-73.
- Hanley, A. B. & Fenwick, G. R. (1985). Cultivated alliums. *J. Plant Foods*, **6**, 211-38.
- Kamanna, V. S. & Chandrasekhara, N. (1984). Hypocholesteremic activity of different fractions of garlic. *Indian J. med. Res.*, **79**, 580-83.
- Kawakishi, S. & Morimitsu, Y. (1988). New inhibitor of platelet aggregation in onion oil. *Lancet*, **ii**, 330.
- Kendler, B. S. (1987). Garlic (*Allium sativum*) and onion (*Allium cepa*): A review of their relationship to cardiovascular disease. *Prev. Med.*, **16**, 670-85.
- Lau, B. H. S., Adetumbi, M. A. & Sanchez, A. (1983). *Allium sativum* (garlic) and atherosclerosis: a review. *Nutr. Res.*, **3**, 119-28.
- Lau, B. H. S., Lam, F. & Wang-Cheng, R. (1987). Effect of an odor-modified garlic preparation on blood lipids. *Nutr. Res.*, **7**, 139-49.
- Liakopoulou-Kyriakides, M., Sinakos, Z. & Kyriakidis, D. A. (1985). Identification of alliin, a constituent of *Allium cepa* with an inhibitory effect on platelet aggregation. *Phytochem.*, **24**, 600.
- Luley, C., Lehmann-Leo, W., Möller, B., Martin, T. & Schwartzkopff, W. (1986). Lack of efficacy of dried garlic in patients with hyperlipoproteinemia. *Arzneim. Forsch./Drug Res.*, **36**, 766-68.
- Lutonski, J. (1984). Klinische Untersuchungen zur therapeutischen Wirksamkeit von Ilja Rogoff Knoblauchpillen mit Rutin. *Z. Phytotherapie*, **5**, 938-42.
- Menon, I. S., Kendal, R. Y., Dewar, H. A. & Newell, D. J. (1968). Effect of onions on blood fibrinolytic activity. *Br. med. J.*, **10 Aug**, 351-52.
- Qureshi, A. A., Din, Z. Z., Abuirmeileh, N., Burger, W. C., Ahmad, Y. & Elson, C. E. (1983). Suppression of avian hepatic lipid metabolism by solvent extracts of garlic: impact on serum lipids. *J. Nutr.*, **113**, 1746-55.
- Sharma, K. K., Gupta, R. K., Gupta, S. & Samuel, K. C. (1977). Antihyperglycemic effect of onion: effect on fasting blood sugar and induced hyperglycemia in man. *Indian J. med. Res.*, **65**, 422-29.
- Sitprija, S., Plengvidhya, C., Kangkaya, V., Bhuvapanich, S. & Tunkayoon, M. (1987). Garlic and diabetes mellitus phase II clinical trial. *J. Med. Ass. Thailand*, **70** (Suppl. 2), 223-27.
- Slater, N. G. P. (1979). Ischaemic heart-disease and garlic. *Lancet*, **i**, 1294.
- Sucur, M. (1980). Effect of garlic on serum lipids and lipoproteins in patients suffering from hyperlipoproteinemia. *Diabetologia Croatica*, **9**, 323-38.
- Üstünes, L., Claeys, M., Laekeman, G., Herman, A. G., Vlietinck, A. J. & Özer, A. (1985). Isolation and identification of two isomeric trihydroxy

- octadenoic acids with prostaglandin E-like activity from onion bulbs (*Allium cepa*). *Prostaglandins*, **29**, 847-65.
- Whitaker, J. R. (1976). Development of flavor, odor, and pungency in onion and garlic. *Adv. Food Res.*, **22**, 73-133.
- Yoshida, S., Kasuga, S., Hayashi, N., Ushiroguchi, T., Matsuura, H. & Nakagawa, S. (1987). Antifungal activity of ajoene derived from garlic. *Applied Environm. Microbiol.*, **53**, 615-17.

(Received 9 May 1989,
accepted 13 July 1989)

CHAPTER 10

CONTROLLED TRIALS IN HUMANS ON THE EFFECTS OF GARLIC SUPPLEMENTS*

Jos Kleijnen

* Submitted for publication

Summary

Deodorized garlic products are said not to have the disadvantages of fresh garlic (odor). The question is whether these commercial garlic supplements are also efficacious. The composition of different garlic preparations probably varies considerably. Consequently, claims about health effects must be substantiated by means of controlled trials for every single garlic supplement.

There are 12 published controlled clinical trials on the efficacy of garlic supplements on cardiovascular risk factors. Besides, there are 6 controlled pharmacological studies with garlic supplements, measuring the effects of a single administered dose on rheological parameters. These trials have been critically assessed, with emphasis on the methodological quality.

The size of the effect found on cholesterol levels and triglycerides is not consistent across the clinical trials: it varies from no effect at all to a difference of almost 20% compared with placebo. However, a majority of the trials report positive results. As for the effects on rheological parameters and blood pressure, differences between the clinical trials are even larger.

In the six pharmacological trials evaluating the effects of a single administration of garlic, it appears that treatment with garlic supplements may result in an increase of blood flow. However, as far as clinical efficacy is concerned, only limited conclusions can be drawn from these small pharmacological studies.

In no trial there has been an attempt of prestratification, and the numbers of patients included are small. Only once, the method of randomization was described, and frequently the methods by which the effects were measured have been insufficiently described. Also, in half of all trials, the data presentation was insufficient to enable the reader to check the assessment of the effects. Most trials have shortcomings on 4 - 6 items of good methodology, and thus there is a serious chance that the results are biased.

Moreover, publication bias complicates the interpretation of reviews and meta-analyses. If there would be a registration of all garlic trials before they are started, all trials - published and unpublished - could be followed up. Also, reviewers could restrict themselves only to registered trials for their main conclusions.

Introduction

Claims have been made on hypoglycaemic effects, antibiotic actions and anticarcinogenic properties of garlic. Convincing evidence about the clinical effects of garlic or garlic products must be obtained by randomized controlled trials in humans. For the above mentioned claims, virtually no evidence from controlled human trials exists. On the other hand, there are controlled trials on the efficacy of garlic in cardiovascular risk factors or risk indicators. Interestingly, claims in this respect include beneficial effects on blood lipids (cholesterol and triglycerides), blood pressure and blood clotting. This is a remarkable combination, because there are not many agents which might have beneficial effects on several risk factors simultaneously. However, cardiovascular risk factors and risk indicators are merely surrogates for the real endpoints in which we are interested: the incidence of cardiovascular disease, cardiovascular mortality and subjective symptoms in patients suffering from cardiovascular diseases.

Controlled trials on the effects of fresh garlic have been reviewed.¹ In these investigations the claims regarding cardiovascular risk factors or risk indicators have been confirmed, but fairly large dosages were used in these experiments: 0.25 - 1 gr/kg body-weight, which is the equivalent of some 7 to 28 cloves of garlic a day. However, the methodological quality of these trials is limited, partly because it is impossible to perform a double blind trial of fresh garlic, but also because there were only small groups of patients, the prognostic comparability of the contrasted groups was not ascertained, and data presentation was far from optimal. In addition to methodological limitations of the garlic trials, the consumption of fresh garlic has disadvantages. Most people do not like the notorious odor, and there are other possible side effects such as gastro-intestinal complaints and allergies.

From deodorized garlic products it is claimed that they do not have these disadvantages, or at least to a lesser extent. The question is whether these commercial garlic supplements are also efficacious, especially because many of the active substances are the same as the ones that cause the odor. Amid 1991, there are 12 published controlled trials on the efficacy of garlic supplements, mostly in patients with elevated levels of blood lipids. We know also of 6 controlled pharmacological studies with garlic supplements, measuring the effects of a single administered dose on rheological parameters.

A major problem in the assessment of the evidence from controlled trials is bias introduced by selective publication. Especially when it concerns small studies, a selection of trials (with positive results?) might have been published, whereas trials reporting negative results might never have been submitted (or accepted) for publication. Publication bias occurs if the results from studies which have not been published, are different from the published ones. It also occurs that only part of the findings is published (for example after multiple testing) or, more subtly, that negative results of a trial are not separately published, but only appear in a meta-analysis from the same author or company, between many other positive findings. In this article, trials will be critically assessed, with emphasis on the methodological quality and with a discussion of the consequences of publication bias.

Materials and methods

Human trials were found by checking references extensively, by personal communication with pharmaceutical companies and by means of a MEDLINE on-line computer search (1966-1991) (keyword: garlic). Trials were only considered if a parallel control group was included.

To explore the possibility that an increasing likelihood of bias (an increasing number of methodological shortcomings) is reflected in the results of the trials, criteria for a methodological assessment of the trials were established in advance. The following criteria have been used for the methodological assessment:

- (A) Patient characteristics adequately described: description of the patients (gender, age, hospitalized or outpatient) and levels of cholesterol, triglycerides, and blood pressure.
- (B) At least 50 patients included per group.
- (C) Prestratification (matching) on relevant prognostic factors.
- (D) Randomization: this was scored if the method of randomization was described and correct.
- (E) Presentation of relevant baseline characteristics.
- (F) Less than 20% dropouts, dropouts described.
- (G) Intervention well described: administration of preparations, dosis and duration of treatment.
- (H) Double blinding.
- (I) Effect measurement relevant and well described: Measurement of the effect must be sensible and reproducible.
- (J) Presentation of the results in such a manner that the analysis can be checked by the reader: depending on the effect measurement, at least the mean(s) and standard deviation, standard error, or confidence interval in each group must be mentioned, or the number of patients with a certain outcome (for example, if rates or proportions were used).

Sometimes only part of the score was given if the description was unclear, or if only some of several interventions, measurements of outcome, or data presentations met the criteria. Concerning the second criterion a fairly large number of participants has been chosen. There are two reasons for this. Firstly, the prognostic comparability at baseline is not guaranteed by randomization of small groups of participants. By chance, prognostically relevant differences might occur. Statistical tests of significance of differences at baseline can be found in reports of many garlic trials, but these are not appropriate because any differences are chance occurrences by definition (there was randomization). Thus, when such differences occur, the biological relevance has to be judged without help of statistics. At the end of the trial, it is often possible to correct for these differences by means of multivariate analysis, but only if the relevant prognostic factors have been measured adequately. The second reason for demanding large numbers of patients is the likelihood of publication bias, which probably decreases as the cost and effort involved increases. If all trials would have been registered beforehand, and if suitable precautions have been taken to assure prognostic comparability at baseline (homogeneous groups of participants, prestratification on important prognostical factors)² the demand for large numbers of participants could be diminished.

Double blinding, when commercial odorless products are involved, should be a lesser problem compared with fresh garlic, and there have also been elegant attempts to make identical placebos by lacing the placebo preparations with a small amount of garlic, so that the active and the placebo medications had a similar garlic smell.³ As for the statistical criterion, we did not demand confidence intervals for the comparisons between groups because then virtually no trials would score the criterion.

Assessment of articles using these criteria provides a score that gives an indication of the methodological quality of each trial. This quality will be an important factor in weighing the conclusions of different trials and, of course, it also should influence the impact on the reader's opinion of all the evidence presented. The function of the total score (table 3) is only a quick reference to the methodological quality, and does not mean that some

sort of numerical equality among the different criteria is presupposed. It could be argued that other criteria should be used for the methodological assessment and that this kind of assessment is rather subjective. However, we selected well established methodological criteria, and the assessment can be checked by the reader (table 3). Also, the reader could use other criteria to see whether substantial changes would occur in the methodological assessment.

Results

There are 12 published controlled trials on the efficacy of garlic supplements in patients on levels of cholesterol and triglycerides, rheological parameters and blood pressure (table 1). In 6 pharmacological studies with garlic supplements, the effects on mainly rheological parameters were measured several hours after the administration of one single dose (table 2). The trials of this last group (with only a single administration of a garlic supplement) cannot be considered as proof of the clinical efficacy of garlic preparations in patients. The methodological quality (table 3), which is aimed at the clinical efficacy, was not assessed for these last 6 trials. In no trial, cardiovascular morbidity or mortality has been assessed, only risk factors or risk indicators were measured. Almost all trials have been performed recently, starting in the second half of the 1980s. In most trials only small numbers of participants have been used. There was only one exception, a trial in which a group of 110 patients on active treatment has been compared with the same number of patients on placebo.¹² Most trials have been performed in Germany, and the duration of the intervention was 4 weeks to 6 months. Compared with the trials of fresh garlic, the dosages which have been used are rather small (equivalent to approx. 2 grams of fresh garlic). This is an interesting fact. The only product of which there is evidence from more than one single trial is Kwai / Sapec, two trade names for essentially the same garlic preparation. Efficacy research is necessary for each individual product, since there are great differences in the fabrication process of the garlic supplements and, consequently, also different concentrations of active substances.

The size of the effect found on cholesterol levels and triglycerides is not consistent across the trials: it varies from no effect at all to a difference of almost 20% compared with placebo. However, a majority of the trials report positive results. For the effects on rheological parameters and blood pressure, differences between the clinical trials are even larger (table 1).

In the six trials evaluating the effects of a single administration of garlic, a decrease has been reported of hematocrit in four studies and a decrease of blood viscosity in three studies (table 2). An increase of capillary blood flow and erythrocyte velocity was reported in 4 studies. Also, in one of the clinical studies, an increase in erythrocyte velocity was reported after 4 weeks of garlic treatment.¹⁴ It appears from these trials that garlic treatment may result in an increase of blood flow. As stated before, with respect to the clinical efficacy of garlic supplements, only limited conclusions can be drawn from these small pharmacological studies.

When the methodological quality of the trials is considered (table 3) it can be concluded that almost all trials have methodological shortcomings. Reading table 3 vertically one immediately sees that shortcomings already occur at the most important moment: the beginning of the experiments. In no trial there has been an attempt of prestratification, and the numbers of patients included are small. Only once, the method of randomization was described, and frequently the methods by which the effects were measured have been insufficiently described. Also, in half of all trials, the data presentation was insufficient to enable the reader to check the assessment of the effects. If the criteria which are used to register medications would be applied to these trials, only the trial of Mader

Table 1

author, year	number	duration	highest daily dosage	cholesterol	triglycerides	rheology	blood pressure	subjective symptoms improved
Lutomski, 1984 ⁴	44/38	12 weeks	300 mg Ilija Rogoff	no change	no change	no change	improvement in 77%/57% of patients	headhache, sleep, dizziness, operating capacity, joy of life, physical well being
Ernst, 1985 ⁵	10/10	4 weeks	600 mg Kwai	-22/-12	-39/-28	no change	no change	nm
Luley, 1986 ⁶	34;51 crossover	6 weeks	600 mg; 1350 mg dried garlic	no change	no change	no change	no change	nm
Lau, 1987 ⁷	16/16	6 months	1000 mg Kyolic	-14/-3	lowering	nm	nm	nm
Sitprija, 1987 ⁸	20/20	1 month	700 mg spray dried garlic	5/1	-3/-5	nm	nm	nm
Barrie, 1987 ³	20 crossover	4 weeks	18 mg cold pressed garlic oil	-8/-2	nm	-16/+6 platelet aggregation	-6/-3	nm
Kandziora, 1988 ⁹	20/20	12 weeks	600 mg Kwai	-6/-2	-8/-1	no change	-14/-14	nm
Kandziora, 1988 ¹⁰	20/20	12 weeks	600 mg Kwai	-10/0	-8/-1	no change	-16/-8	nm
Auer, 1989 ¹¹	24/23	12 weeks	600 mg Kwai	-14/-7	-18/-6	nm	-13/-4	nm
Mader, 1990 ¹²	110/111	16 weeks	800 mg Sapec	-12/-3	-17/-2	nm	positive trend for garlic group	lack of energy, fatigue
Vorberg, 1990 ¹³	20/20	16 weeks	900 mg Sapec	-21/-3	-24/-5	nm	only small change	activeness, energy, good mood, fatigue, anxiety, sensitivity, agitation, irritability
Kiesewetter, 1991 ¹⁴	30/30	4 weeks	800 mg Kwai	no change	no change	Decreased platelet aggregation. Erythrocyte velocity in nail fold capillaries: +49/0	insufficient data presented	nm

"/" divides garlic group from placebo group; nm = not measured

might be judged to be of acceptable quality for the assessment of the efficacy of a garlic preparation. Most trials have shortcomings on 4 - 6 items of good methodology, and thus there is a serious chance that the result are biased.

The trial with the best methodological quality has been performed by the study group on phytotherapy of the German Association of General Practitioners. This was a multicenter placebo-controlled randomized study with parallel groups conducted in 30 practices of general practitioners in Germany.¹² Patients received either 4 garlic tablets containing a total of 800mg garlic powder, which was standardized to 1.3% alliin, or placebo tablets which were identical in appearance, for a period of 4 months. A total of 261 patients

Table 2

author, year	number	duration	highest daily dosage	cholesterol	triglycerides	rheology	blood pressure	subjective symptoms improved
Jung, 1989 ¹⁵	10/10	once	1200 mg Kwai	nm	nm	Hematocrit: -6/0, viscosity: -6/0. No change for several other parameters	no change	nm
Kiesewetter, 1990 ¹⁶	10 crossover	once	600 mg Kwai	nm	nm	Hematocrit: -3/0, viscosity: -2/0, fibrinolytic activity (t-PA): +51/+35	nm	nm
Jung, 1990 ¹⁷	10 crossover	once	900 mg Kwai	nm	nm	Erythrocyte velocity in skin capillaries: +55/+10 Cutaneous capillary blood flow: +45/+5.	+4/-3	nm
Wolf, 1990 ¹⁸	20 crossover	once	900 mg Kwai	nm	nm	Diameter conjunctival vessels. Arterioles: +4/0 Venules: +6/-1 Capillaries: 0/0	nm	nm
Jung, 1991 ¹⁹	10 crossover	once	800 mg +	no change	no differences	Hematocrit: -3/+2 Viscosity: +2/+3 Erythrocyte velocity: +4/-20	nm	nm
Koscielny, 1991 ²⁰	8 crossover	once	1. 900 mg Kwai 2. 270 mg Sanhelios 3. 5 mg Xund	nm	nm	Hematocrit: 1. -3/0 2. -2/0 3. -2/0 Viscosity: 1. -3/+2 2. +1/+2 3. 0/+2 Erythrocyte velocity: 1. +27/-13 2. -1/-13 3. +15/-13	no change	nm

"/" divides garlic group from placebo group; nm = not measured

with cholesterol levels of 200-300 mg/dl and/or triglyceride values of 200-300 mg/dl entered the trial. Relevant baseline characteristics were presented in detail for the active and placebo groups. There was no prestratification, but the method of randomization was described (numbers generated on a computer and balanced in blocks of 10 per investigator). Fifteen percent of the randomized patients dropped out of the study, and both the reasons and the distribution between the groups of dropouts and participants were described. Cholesterol levels in the active group dropped from 266 to 235 mg/dl compared with a drop from 262 to 255 mg/dl in the placebo group. In the active group, triglyceride values dropped from 226 to 188 mg/dl and in the placebo group from 217 to

213 mg/dl. There was no check of compliance. However, patients were asked whether they or others had been aware of any garlic odor. Garlic smell was mentioned in up to 21% of the cases in the active group, and also in up to 9% of the placebo group. This indicates the necessity of checking whether the trial was indeed double blind. In this case, blinding was partially broken, a fact which might have impact on the results. However, it is very difficult to estimate the magnitude of this bias. Of course, it is interesting to assess the results for the patients who are still blinded in each group. However, such an analysis ignores the fact that also the still blinded people from both groups may not be comparable anymore. Finally, it must be mentioned that in all other trials there was no check of blinding at all, and critics of garlic therapy could very well assert that the results of all these trials are potentially (or even probably) biased.

Table 3. The methodological score assigned to the experiments

first author, year	A	B	C	D	E	F	G	H	I	J	TOT
Lutomski, '84 ⁴	±	—	—	±	—	+	+	+	±	±	5
Ernst, '85 ⁵	±	—	—	±	—	—	+	—	±	—	2.5
Luley, '86 ⁶	—	—	—	±	—	+	+	+	±	—	4
Lau, '87 ⁷	±	—	—	±	±	+	+	+	±	±	5.5
Sitprijja, '87 ⁸	+	—	—	±	+	±	±	+	±	+	6
Barrie, '87 ³	±	—	—	±	+	+	+	+	±	+	6.5
Kandziora, '88 ⁹	+	—	—	±	+	+	—	—	+	+	5.5
Kandziora, '88 ¹⁰	±	—	—	±	+	+	+	+	+	±	6.5
Auer, '89 ¹¹	+	—	—	±	+	—	+	+	±	+	6
Mader, '90 ¹²	+	+	—	+	+	+	+	+	±	+	8.5
Vorberg, '90 ¹³	+	—	—	±	+	+	+	+	±	+	7
Kiesewetter, '91 ¹⁴	±	—	—	±	—	—	+	+	+	±	4.5

+ : item is scored

— : item is not scored

± : item is partly scored (25-50 patients in category B)

TOT : total number of items which were scored

A: Patient characteristics adequately described

B: At least 50 patients included per group

C: Prestratification (matching) on relevant prognostic factors

D: Randomization

E: Presentation of relevant baseline characteristics

F: Less than 20% dropouts, dropouts described

G: Intervention well described

H: Double blinding

I: Effect measurement relevant and well described

J: Presentation of the results

Discussion

The composition of different garlic preparations probably varies considerably. It might well be that after processing and storage, a part or even all of the active ingredients have disappeared. The contents of the preparations are difficult to establish, and the consumer must already be contended if the product is at least standardized on one or two (active?) ingredients. Consequently, claims about health effects must be substantiated by means of controlled trials for every single garlic supplement. Controlled trials on the efficacy of single active components of garlic are completely lacking, although it is pos-

sible to make several of the substances synthetically which are claimed to have pharmacological actions.

There is an interesting contrast between the dosages used in trials of fresh garlic (many grams a day) and in most trials of garlic supplements (equivalent to 1-2 grams of fresh garlic a day). Of course, for supplements, low dosages have been chosen because of the odor, since the active ingredients are the same ingredients which cause the odor. Are the dosages which have been used in trials of fresh garlic too high, or must we be careful when there are claims of supplements with lower dosages? It might also be that in order to cause changes in blood lipid levels and blood pressure, higher dosages are needed than those needed to influence for example rheological parameters. The existing evidence from controlled trials in humans is inconclusive to answer these questions.

For future trials on the efficacy of garlic supplements, the contrasted groups must be prognostically comparable at baseline. To ensure this, there are several strategies available: restriction to more homogeneous groups of patients, prestratification, using large numbers of participants (if possible), random allocation, and after randomization presentation of baseline characteristics for active and placebo groups. The method of fabrication of the substance which is used must be described, and it must be mentioned on which active ingredients it has been standardized. The study of Mader has shown how important it is to check double blinding. For the blinded effect measurement relevant outcome parameters must be used, which would include not only cardiovascular risk factors / risk indicators, but also the effects on subjective complaints which might be manifestations of underlying cardiovascular disease. Well performed trials comparing the effects of garlic supplements with registered products for similar indications are lacking. To be truly convincing, also the effects on cardiovascular morbidity / mortality should be assessed. However, such trials would cost an enormous amount of money, and they would take many participants who must be willing to be on placebo treatment for a long time. Whenever possible, raw data with means and confidence intervals should be presented, not only for the separate groups, but also for the between group comparisons. The value of review articles depends, among other factors, on how completely and how validly the included publications represent the total available evidence. A major problem in summarizing the existing literature is finding all publications addressing the subject of interest. However, even if all the published literature is identified and included in a review, the problem remains that there might be relevant research findings that have not been published. Publication bias complicates the interpretation of reviews and meta-analyses. If favorable results are published more often, there will be an overestimation of the effects of a treatment. Although it is probable that part of all performed garlic studies has not been published, it is impossible to judge the extent of publication bias in research of garlic supplements, and the influence on the existing state of affairs. How can we make sure that publication bias does not occur? If there would be a registration of all garlic trials before they are started, all trials - published and unpublished - could be followed up. Also, reviewers could restrict themselves to registered trials for their main conclusions. Other information could then be viewed as most of nowadays reviews: prone to publication bias.

Conclusions

There are 12 published clinical trials on the efficacy of garlic supplements on cardiovascular risk factors. There are also 6 controlled pharmacological studies with garlic supplements, measuring the effects of a single administered dose on rheological parameters. The size of the effect found on cholesterol levels and triglycerides is not consistent across the clinical trials: it varies from no effect at all to a difference of almost 20% compared

with placebo. However, a majority of the trials report positive results. For the effects on rheological parameters and blood pressure, differences between the clinical trials are even larger. In the six pharmacological trials evaluating the effects of a single administration of garlic, it appears that treatment with garlic supplements may result in an increase of blood flow.

In half of all trials, the data presentation was insufficient to enable the reader to check the assessment of the effects. Most trials have shortcomings on 4 - 6 items of good methodology, and thus there is a serious chance that the result are biased. The interpretation of reviews and meta-analyses is further complicated by publication bias.

References

1. Kleijnen J, ter Riet G, Knipschild P. 1989. Garlic, onions and cardiovascular risk factors. A review of the evidence from human experiments with emphasis on commercially available preparations. *Br J Clin Pharmacol* 28: 535-544.
2. Knipschild PG, Leffers P, Feinstein AR. 1991. The qualification period. *J Clin Epidemiol* 44: 461-464.
3. Barrie SA, Wright JV, Pizzorno JE. 1987. Effects of garlic oil on platelet aggregation, serum lipids and blood pressure in humans. *J Orthomol Med* 2: 15-21.
4. Lutomski J. 1984. Klinische Untersuchungen zur therapeutischen Wirksamkeit von Ilja Rogoff Knoblauchpillen mit Rutin. *Z Phytotherapie* 5: 938-942.
5. Ernst E, Weihmayr T, Matrai A. 1985. Garlic and blood lipids. *Br Med J* 291: 139.
6. Luley C, Lehmann-Leo W, Möller B, Martin T, Schwartzkopff W. 1986. Lack of Efficacy of dried garlic in patients with hyperlipoproteinemia. *Arzneim Forsch/Drug Res* 36: 766-768.
7. Lau BHS, Lam F, Wang-Cheng R. 1987. Effect of an odor-modified garlic preparation on blood lipids. *Nutrition Res* 7: 139-149.
8. Sitprija S, Plengvidhya C, Kangkaya V, Bhuvapanich S, Tunkayoon M. 1987. Garlic and diabetes mellitus phase II clinical trial. *J Med Ass Thailand* 70 (suppl 2): 223-227.
9. Kandizora J. 1988. Antihypertensive Wirksamkeit und Verträglichkeit eines Knoblauch-Präparates. *Ärztliche Forschung* 35(1): 1-8.
10. Kandizora J. 1988. Blutdruck- und lipidsenkende Wirkung eines Knoblauch-Präparates in Kombination mit einem Diuretikum. *Ärztliche Forschung* 35(3):1-8.
11. Auer W, Eiber E, Hertkorn E, et al. 1989. Hypertonie und Hyperlipidämie: In leichteren Fällen hilft auch Knoblauch. Multizentrische placebokontrollierte Doppelblind-Studie zur lipid- und blutdrucksenkenden Wirkung eines Knoblauchpräparates. *Der Allgemeinarzt* 11: 205-208.
12. Mader FH. 1990. Treatment of hyperlipidaemia with garlic-powder tablets. Evidence from the German Association of General Practitioners' multicentric placebo-controlled double-blind study. *Arzneimittelforschung* 40: 1111-1116.
13. Vorberg G, Schneider B. 1990. Therapie mit Knoblauch. Ergebnisse einer placebokontrollierten Doppelblindstudie. *Natur und Ganzheitsmedizin* 3: 62-66.
14. Kiesewetter H, Jung F, Pindur G, et al. 1991. Effect of garlic on thrombocyte aggregation, microcirculation, and other risk factors. *Int J Clin Pharmacol Ther Toxicol* 29(4): 151-155.
15. Jung F, Kiesewetter, Mrowietz C, et al. 1989. Akutwirkung eines zusammengesetzten Knoblauchpräparates auf die Fließfähigkeit des Blutes. *Zschr Phytotherapie* 10: 87-91.
16. Kiesewetter H, Jung F, Mrowietz C, et al. 1990. Effects of garlic on blood fluidity and fibrinolytic activity: a randomised, placebo-controlled, double-blind study. *Br J Clin Pract* 44(Suppl 69): 24-29.
17. Jung F, Jung EM, Mrowietz C, et al. 1990. Influence of garlic powder on cutaneous microcirculation: a randomised, placebo-controlled, double-blind, crossover study in apparently healthy subjects. *Br J Clin Pract* 44(Suppl 69): 30-35.
18. Wolf S, Reim M. 1990. Effect of garlic on conjunctival vessels: a randomised, placebo-controlled, double-blind trial. *Br J Clin Pract* 44(Suppl 69): 36-39.
19. Jung F, Kiesewetter H. 1991. Einfluß einer Fettbelastung auf Plasmalipide und die Fließfähigkeit des Blutes sowie Senkung der Hämatokritwerte in Abhängigkeit von der Dosis. *Medizinische Welt* 42(-Suppl 7A): 14-17.
20. Koscielny J, Jung EM, Jung F, et al. 1991. Einfluß von Knoblauchpräparate auf die Fließfähigkeit des Blutes sowie die periphere Mikrozirkulation. Bei Knoblauchtrockenpulver signifikante Wirkung festgestellt. *Medizinische Welt* 42(Suppl 7A): 29-31.

CHAPTER 11

GINKGO BILOBA FOR INTERMITTENT CLAUDICATION AND CEREBRAL INSUFFICIENCY*

Jos Kleijnen, Paul Knipschild

* Submitted for publication

Summary

Objective - To establish whether there is evidence from controlled trials in humans on the efficacy of Ginkgo biloba extracts in intermittent claudication and cerebral insufficiency.

Design - Criteria-based meta-analysis. Assessment of the methodological quality of 15 trials on intermittent claudication and 38 trials on cerebral insufficiency. Trials were scored using a list of predefined criteria of good methodology, and the outcome of the trials was interpreted in relation to their quality. A comparison of the quality was made with trials of pentoxifylline and co-dergocrine, which are registered products for the same indications.

Setting - Controlled trials published in any language.

Main outcome measures - Results of the trials with the best methodological quality. Trials in intermittent claudication and in cerebral insufficiency were considered separately.

Results - For intermittent claudication only 2 trials were reasonably well performed; for cerebral insufficiency there were 8 well performed trials. Shortcomings were limited numbers of patients included, and incomplete description of randomization procedures, patient characteristics, effect measurement and data presentation. In no trial double blindness was checked. For both indications virtually all trials reported positive results, in most trials the dosage was 120 mg Ginkgo extract a day, given for at least 4 - 6 weeks. For the best trials, there were no marked differences in the quality of the evidence of the efficacy of Ginkgo in cerebral insufficiency compared with co-dergocrine. The results of the review may be complicated by a combination of publication bias and other biases, because there were no negative results reported in many trials of low methodological quality.

Conclusions - Positive results have been reported for Ginkgo biloba extracts in the treatment of cerebral insufficiency and intermittent claudication. The clinical evidence is similar to that of some widely registered products which are prescribed for the same indications. However, further studies should be conducted for a more detailed assessment of the efficacy.

Introduction

Extracts from the leaves of the Ginkgo biloba tree (maidenhair tree) are used by patients with complaints of peripheral vascular disease such as intermittent claudication, and in the treatment of cerebral insufficiency. Symptoms in elderly people such as dizziness, tinnitus, headaches, difficulties of concentration and memory, lack of energy, depressive mood and anxiety are claimed to be relieved by treatment with Ginkgo. These symptoms have been associated with decreased cerebral circulation; sometimes they are considered early symptoms of dementia either of the degenerative or the multiple infarctions type. Often no explanation for the symptoms is found.

Several mechanisms of action from Ginkgo extracts have been described: (1) Effects on blood circulation such as vasoregulating activity of arteries, capillaries, veins (increased blood flow) and rheological effects (decreased viscosity, antagonistic to platelet activating factor receptors); (2) Metabolic changes, for example on neuron metabolism (increased tolerance for anoxia); (3) Beneficial influence on neurotransmitter disturbances; (4) Prevention of damage of membranes caused by free radicals. These effects may be caused by single active ingredients, or by the combined action of the many active agents found in Ginkgo extracts. The most important are considered flavonoid substances (Ginkgo-flavone glycosides, biflavonoids) and terpenoids (ginkgolides and bilobalide).^{1,2,3,4} There are 4 Ginkgo preparations which have been used in controlled trials: Tebonin®, Tanakan®, rōkan® and Kaveri®. The first three are different names for the same extract called EGB 761, and they are standardized on the amount of Ginkgo-flavone glycosides (24%) and terpenoids (6%). The fourth is standardised on the same ingredients in comparable dosages (25% Ginkgo-flavone glycosides, and recently also 6% terpenoids). The dosage is 3 tablets of 40 mg Ginkgo extract a day, given for at least 4-6 weeks.

In Germany, Ginkgo biloba is one of the most frequently prescribed drugs: more than 5 million prescriptions in 1988, at a cost for the health insurance of 370 million DM.^{5,6} In 1989, this cost was for Tebonin alone 219 million DM, and for rōkan 62 million DM. This is by far the highest amount, followed by Adalat (nifedipine) with 195 million DM.⁷ In France there is also a large Ginkgo market, whereas in other countries such as the Netherlands, the U.K. and the U.S.A. most people and physicians have never heard of Ginkgo.

In this article, the evidence from controlled trials in humans of the clinical efficacy of Ginkgo biloba will be presented. The methodological quality of these trials will be emphasized. Almost all trials have been done in patients with cerebral insufficiency and patients with intermittent claudication. Because in many countries Ginkgo is not registered, we also compared the quality of the evidence with that of two widely registered products for similar indications: pentoxifylline (Trental®) for intermittent claudication and ergoloid mesylates, also called co-dergocrine (Hydergine®) for cerebral insufficiency.

Materials and methods

Trials in humans assessing the effects of Ginkgo on intermittent claudication and clinical symptoms of cerebral insufficiency were eligible if parallel index and control groups were included. Crossover trials were also eligible. Trials in healthy volunteers were excluded.

Experiments were found by means of various strategies:

(1) computer searches: MEDLINE online 1966-91 (Keyword: Ginkgo biloba) and EM-BASE online 1974-91 (Keywords: Ginkgo, Ginkgo biloba); (2) checking references extensively, in articles on clinical research and in textbooks;⁸ (3) personal communication with researchers; (4) correspondence with and visiting major manufacturers of Ginkgo biloba

medications. Trials published in any language were eligible, without restrictions. Interestingly, two trials we knew of were not identified in our computer search.^{9,10} When we checked, it turned out that these publications can be found using the free text terms *Ginkgo* and *Gingko*. We identified several other articles on Ginkgo in which the word was spelled as Gingko! It could well be that in other subjects this problem of computer searches will also occur.

We asked the manufacturers of Trental® and Hydergine® to send us the best clinical trials of their product. We compare the best studies of Ginkgo biloba with 5 studies of each product.

To explore the possibility that an increasing number of methodological shortcomings is reflected in the results of the trials, criteria for a methodological assessment of the experiments were established. A motivation and discussion of these criteria and scores has been published earlier.¹¹

Starting from a maximum score of 100 points, we divided these among 7 methodological criteria.

- (A) Patient characteristics adequately described: 10 points. - Description of the patients (gender, age, hospitalized or outpatient) and the symptoms, including their duration and severity. If appropriate, also a description of relevant prognostic factors (e.g. walking distance, diabetes and smoking habits for intermittent claudication).
- (B) Number of patients analysed: 30 points. - 100 or more patients per group = 30 points, 75-99 patients per group = 25 points, 50-74 patients = 20 points, 40-50 patients = 15 points, 25-40 patients = 10 points. A crossover trial with 70 participants (35 given active treatment and 35 placebo in each period) would score 10 points.
- (C) Randomization: 20 points. - Twenty points if the method of randomization was described and correct, 10 points if the method was not described or if some form of pseudo-randomization was applied. If there were fewer than 25 participants per group, half the score was given unless there was prestratification (matching) on relevant items and a table showing comparable baseline characteristics.
- (D) Intervention well described: 5 points. - Administration of Ginkgo preparations: dosis and duration of treatment.
- (E) Double blinding: 20 points. - Twenty points if the placebo was described as indistinguishable, 10 points if double blinding was only mentioned.
- (F) Effect measurement relevant and well described: 10 points. - Measurement of the effect must be sensible and reproducible. Five points each for relevance and adequate description.
- (G) Presentation of the results in such a manner that the analysis can be checked by the reader: 5 points. - Depending on the effect measurement, at least the mean(s) and standard deviation, standard error, or confidence interval in each group must be mentioned, or the number of patients with a certain outcome (for example, if rates or proportions were used).

Sometimes only part of the score was given if the description was unclear, or if only some of several interventions, measurements of outcome, or data presentations met the criteria. In the second criterion we chose to use the number of patients analysed instead of the number randomized because in many publications dropouts were not accounted for. For the effect measurement, because we wanted to know whether Ginkgo treatment reduces the patient's complaints, we considered clinical effects (severity of symptoms) more important than pathophysiological effects, such as ankle/arm blood pressure index or an electroencephalogram. As for the statistical criterion, we did not demand confidence intervals for the comparisons between groups because then virtually no trials would score the criterion.

All articles were scored at least twice, and differences mostly caused by reading errors were resolved by discussions. The largest difference was 15 points. Assessment of articles using these criteria provides a score that gives an indication of the methodological quality of each trial.

Results

Intermittent claudication

From a total of 15 studies for this indication, 8 have been performed in Germany, 5 in France 1 in Italy and 1 in the U.K. Three of these were published in journals indexed in the Index Medicus. Only two trials are of acceptable quality (75 and 66 points in table 1), but they are certainly not perfect. The shortcomings can easily be detected reading table 1 vertically. The most frequent shortcomings are small numbers of patients and inadequately described randomization. Also, the description of patient characteristics and effect measurement, and the data presentation were often insufficient. However, the results are consistent: all trials indicated positive effects; we were unable to find any negative trial.

Table 1. Intermittent claudication

author, year	preparation	result	CH 10	NU 30	RA 20	IN 5	DB 20	EF 10	DA 5	TOT 100
Bauer, '84 ^{12,13}	Rökan	sign.	10	15	10	5	20	10	5	75
Saundreau, '89 ¹⁴	Tanakan	sign.	8	10	10	5	20	10	3	66
Bulling, '91 ¹⁵	Rökan	sign.	8	0	10	5	20	8	5	56
Salz, '80 ¹⁶	Rökan	sign.	6	0	10	5	20	10	0	51
Thomson, '90 ¹⁰	Tanakan	pos.	5	0	5	5	20	10	3	48
Garzya, '81 ¹⁷	Tanakan	pos.	6	0	5	5	20	10	0	46
Courbier, '77 ¹⁸	Tanakan	sign.	8	0	0	5	20	8	0	41
Böhmer, '88 ¹⁹	Rökan	pos.	10	0	5	5	10	10	0	40
Rudofsky, '87 ²⁰	Tebonin	pos.	4	0	5	5	10	8	5	37
Frileux, '75 ²¹	Tanakan	sign.	4	0	0	5	20	3	3	35
Diehm, '90 ²²	Tebonin	sign.	4	0	5	5	10	5	0	29
Ambrosi, '75 ²³	Tanakan	sign.	6	0	0	5	10	3	3	27
Baitsch, '86 ²⁴	Rökan	pos.	8	0	5	5	0	5	3	26
Berndt, '87 ²⁵	Rökan	sign.	8	0	5	5	0	8	0	26
Bastide, '78 ²⁶	Tanakan	pos.	6	0	0	3	0	5	0	14

Legend:

TOT : total number of points scored

sign. : index group fared better (author stated that result was significant)

pos. : positive trend (or significant effect for only some of the effect measurements)

CH: Patient characteristics; NU: Number of patients per group; RA: Randomization; IN: Intervention well described; DB: Double blind; EF: Effect measurement well described; DA: Data presentation.

The trial of Bauer,^{12,13} in patients with Fontaine's stage IIb claudication (pain-free walking distance less than 100 m), shows an average increase of the total distance walked from 112 to 222 meters on Ginkgo compared with an increase from 145 to 176 meters on placebo. The treatment, 120 mg Ginkgo extract daily, was given for 6 months. The trial of Saundreau et al,¹⁴ in patients with pain in rest (stage III of Fontaine's classification), showed a decrease on a 100 mm visual analogue scale for pain from 61 to 30 mm on

infusions of 200 mg Ginkgo extract for 8 days, compared with a decrease from 51 to 39 mm in the placebo group. The other trials showed similar or even greater positive effects.

Cerebral insufficiency

In Germany, 20 studies have been performed for cerebral insufficiency; 14 studies have been performed in France, 3 studies come from Italy and 1 from the U.K.. Fourteen of these 38 studies were published in journals indexed in the Index Medicus. Compared with intermittent claudication there is more and better evidence for the efficacy of Ginkgo in cerebral insufficiency (table 2). Eight trials are of acceptable quality but, again, not perfect. Six of these are briefly described below. For the other trials, the methodological assessment shows the same shortcomings as in intermittent claudication: small numbers of patients, inadequate description of randomization procedures, patient characteristics, effect measurement and data presentation. As before, all trials show positive effects compared with placebo on symptoms as described in the introduction part. The only trial with negative results was in patients with dementia of vascular origin.⁵⁶

Schmidt et al²⁷ treated 99 out-patients (average age 59) with the diagnosis "cerebral insufficiency" for 12 weeks, with Ginkgo extract containing a daily dose of 150 mg Ginkgo-flavone glycosides. Fifty patients received Ginkgo and 49 placebo. There were no dropouts. Positive effects were found on difficulties with concentration and memory, tiredness, operative capacity and other symptoms. Brüchert et al, in a multi-center trial of the German Association of General Practitioners, assessed the efficacy of 150 mg Ginkgo extract for 12 weeks in 303 out-patients with cerebral insufficiency.²⁸ The average age was 69 and the average duration of the complaints was 46 months. There were 94 dropouts. Reasons for exclusion from the analysis were effect measurement at the wrong time in 40 cases (verum/placebo: 21/19), inclusion criteria not met (11/12), concomitant treatments which were not allowed (10/15), and discontinuation of treatment (4/11). One hundred ten patients received Ginkgo and 99 placebo. Positive effects were found on difficulties with concentration and memory, tiredness, operative capacity and other symptoms. Meyer²⁹ treated patients (average age 50) with tinnitus and associated symptoms, such as dizziness and hearing impairment. The average duration of the complaints was 4-5 months; Ginkgo (4 ml containing 160 mg Ginkgo extract a day) or placebo was given for 3 months; and patients were followed up for 13 months. Improvement or cure was found after an average of 70 days in patients receiving Ginkgo compared with 119 days in patients treated with placebo. Taillandier et al³⁰ in elderly patients (average age 82) with symptoms such as dizziness, tinnitus, headaches and difficulties of concentration and memory found positive effects after three months' daily treatment with 160 mg Ginkgo extract. For the effect measurement, they used a scale for the clinical assessment of geriatric patients (l'échelle d'appréciation clinique en gériatrie, E.A.C.G.). Haguenaer et al³¹ assessed the effects in patients (average age 50) with a vertiginous syndrome of recent onset and undetermined origin. After 3 months, symptoms disappeared in 47% of patients treated daily with 160 mg Ginkgo extract compared with 18% of those on placebo. Vorberg et al³² gave daily 112 mg Ginkgo extract, containing 30 mg Ginkgo-flavone glycosides to patients with an average age of 70 years. After 12 weeks' treatment positive effects were found for difficulties of concentration and memory, anxiety, dizziness and headaches.

Table 2. Cerebral insufficiency

author, year Maximum	preparation	result	CH 10	NU 30	RA 20	IN 5	DB 20	EF 10	DA 5	TOT 100
Schmidt, '91 ²⁷	Kaveri	sign.	10	20	20	5	20	10	5	90
Brüchert, '91 ²⁸	Kaveri	sign.	10	30	10	5	10	10	5	80
Meyer, '86 ²⁹	Tanakan	sign.	10	20	10	5	20	10	5	80
Taillandier, '86 ³⁰	Tanakan	sign.	10	25	10	5	20	10	0	80
Haguenaue, '86 ³¹	Tanakan	sign.	10	10	20	5	20	10	3	78
Vorberg, '89 ³²	Kaveri	sign.	10	15	10	5	20	10	5	75
Eckmann, '90 ³³	Kaveri	sign.	10	10	10	5	20	7	5	67
Wesnes, '87 ³⁴	Tanakan	sign.	7	10	10	5	20	10	5	67
Augustin, '76 ³⁵	Tanakan	sign.	6	25	10	5	10	5	0	61
Arrigo, '86 ^{36,37}	Tebonin	sign.	9	15	10	5	10	10	3	61
Halama, '91 ³⁸	Kaveri	sign.	10	0	10	5	20	10	5	60
Meyer, '86 ³⁹	Tanakan	sign.	10	25	10	0	0	10	5	60
Halama, '88 ⁹	Tebonin	sign.	8	0	10	5	20	10	3	56
Chesseboeuf, '79 ⁴⁰	Tanakan	pos.	10	10	20	5	0	5	5	55
Maier-Hauff, '91 ⁴¹	Kaveri	sign.	10	0	10	5	20	8	0	53
Weitbrecht, '86 ^{42,43}	Tebonin	sign.	10	0	5	5	20	10	3	53
Schwerdtfeger, '81 ⁴⁴	Rökan	sign.	8	0	5	5	20	10	5	53
Moreau, '75 ⁴⁵	Tanakan	sign.	3	10	10	5	20	3	0	51
Dieli, '81 ⁴⁶	Tanakan	sign.	7	0	5	5	20	8	5	50
Eckmann, '82 ⁴⁷	Tebonin	sign.	5	10	10	5	10	5	5	50
Dubreuil, '86 ⁴⁸	Tanakan	pos.	9	0	20	5	5	10	0	49
Hofferberth, '91 ⁴⁹	Kaveri	sign.	5	10	10	5	10	8	0	48
Hofferberth, '91 ⁵⁰	Rökan	sign.	6	0	5	5	20	5	5	46
Hofferberth, '89 ⁵¹	Rökan	sign.	5	0	5	5	20	5	5	45
Israël, '87 ⁵²	Tanakan	pos.	6	0	5	5	20	8	0	44
Haan, '82 ⁵³	Tebonin	pos.	6	10	10	5	0	8	5	44
Claussen, '84 ⁵⁴	Rökan	sign.	6	0	5	5	20	5	3	44
Gerhardt, '90 ⁵⁵	?	pos.	6	10	10	5	0	10	0	41
Hartmann, '91 ⁵⁶	Kaveri	neg.	6	10	0	5	10	8	0	39
Geßner, '85 ⁵⁷	Rökan	pos.	4	0	5	5	10	8	5	37
Agnoli, '80 ⁵⁸	Tebonin	sign.	6	0	5	5	10	10	0	36
Hamann, '85 ⁵⁹	Tebonin	pos.	8	0	5	5	10	5	3	36
Teigeler, '84 ⁶⁰	Tebonin	pos.	3	0	5	5	10	3	5	31
Pidoux, '83 ⁶¹	Tanakan	sign.	3	0	5	5	10	5	0	28
Chartres, '87 ⁶²	Tanakan	pos.	8	0	5	0	10	5	0	28
Natali, '79 ⁶³	Tanakan	pos.	7	0	0	5	0	10	5	27
Bono, '75 ⁶⁴	Tanakan	sign.	4	0	0	5	10	5	0	24
Israël, '77 ⁶⁵	Tanakan	pos.	3	0	0	5	0	5	5	18

Legend:

TOT : total number of points scored

sign. : index group fared better (author stated that result was significant)

pos. : positive trend (or significant effect for only some of the effect measurements)

neg. : no difference between Ginkgo and placebo groups

CH: Patient characteristics; NU: Number of patients per group; RA: Randomization; IN: Intervention well described; DB: Double blind; EF: Effect measurement well described; DA: Data presentation.

No serious side effects were reported in any trial and, if present, side effects were not different from those in patients treated with placebo. In most trials, the dosage used is 120 - 160 mg a day; whether other dosages have different effects remains to be determined. From most trials it appears that the treatment should be given for at least 4 to 6

weeks before positive effects can be expected. However, in one well performed trial positive effects were reported after one week treatment, with a higher dosage of 200 mg a day.¹⁴

Published trials of pentoxifylline and co-dergocrine which we obtained from the pharmaceutical firms had equal or slightly higher scores on our methodological assessment. Table 3 shows the assessment of five trials of each product which according to their firms were well performed. A comparison shows that the quality of the trials on the efficacy of these registered products is not much different compared with the best evidence on Ginkgo available at the moment. A direct comparison of Ginkgo and pentoxifylline was made by Böhmer et al¹⁹ in Fontaine stage IIb claudicatio patients. Treatment with daily 160 mg Ginkgo extract or 1200 mg pentoxifylline during 24 weeks showed similar improvements (200-250 meters) of the walking distance for both treatments. A direct comparison of daily 120 mg Ginkgo and 4.5 mg dihydroergotoxine was made by Gerhardt et al⁵⁵ in patients (55-85 years) with symptoms such as dizziness, difficulties of memory and concentration, headaches and depressive mood. After 6 weeks both groups showed improvements, comparisons between groups showed no differences.

Table 3.

author, year	preparation	result	CH	NU	RA	IN	DB	EF	DA	TOT
Maximum			10	30	20	5	20	10	5	100
Trental & Intermittent Claudication:										
Lindgärde '89 ⁶⁶	Trental	sign.	10	25	10	5	20	10	5	85
Tonak '77 ⁶⁷	Trental	sign.	9	20	10	5	20	10	3	77
Porter '82 ⁶⁸	Trental	sign.	9	15	10	5	20	10	5	74
Kellner '83 ⁶⁹	Trental	sign.	9	0	5	5	20	10	5	54
Roekaerts '84 ⁷⁰	Trental	sign.	9	0	5	5	20	10	5	54
Hydergine & Mental Functioning:										
Lazzari '83 ⁷¹	Hydergine	sign.	9	30	10	5	20	10	0	84
Yoshikawa '83 ⁷²	Hydergine	sign.	8	30	10	5	20	5	5	83
Bargheon '73 ⁷³	Hydergine	sign.	5	20	20	5	20	10	0	80
McConnachie '73 ⁷⁴	Hydergine	sign.	9	10	20	5	20	10	5	79
Hollingsworth '80 ⁷⁵	Hydergine	sign.	9	10	20	5	20	10	0	74

Legend:

TOT : total number of points scored

sign. : index group fared better (author stated that result was significant)

CH: Patient characteristics; NU: Number of patients per group; RA: Randomization; IN: Intervention well described; DB: Double blind; EF: Effect measurement well described; DA: Data presentation.

Discussion

There are many controlled trials on the efficacy of Ginkgo, 15 in patients with intermittent claudication and 38 in patients with cerebral insufficiency. The authors of nearly all trials report at least a partially positive outcome, but sometimes we had difficulties to see this in the data presented.^{8,10} Considering the low quality of many trials, we find it hard to believe that there are no small trials with negative results. Probably such trials exist, but have not been published. On the other hand, large well-performed trials, which have

taken much costs and effort, will be published regardless of the outcome. Because of this, we will base our conclusions only on these well-performed trials.

A problem with our methodological assessment is that limited description of the methods and the results in the publication may lead to a lower score. We believe, however, that a detailed description of this information is as important as using good methodology in practice. It would surely be very complicated if every single reader must go after the necessary information. Also, we scored the trials from the viewpoint of the practitioner who treats patients. From this point of view, information about what happens to the patient's symptoms is more important than for instance a change of the EEG, or the influence on blood rheology.

It could be argued that other criteria should be used for the methodological assessment and that this kind of assessment is rather subjective. However, we have selected well established criteria. Moreover, we see the total score only as an indication whether a trial is well performed or not. The scores on the individual items may give the reader an impression why we arrived at such a conclusion. We would like to emphasize that scoring five points more or less is not very important to us. Furthermore, it is always possible that a well performed trial has one fatal mistake, for instance a break of double blindness, which would render the results worthless. The probability for making such a mistake may increase when the number of methodological shortcomings increases, i.e. when the score decreases. Unfortunately, we were not blinded in assessing the methodological score of every publication. We were already too familiar with the publications before we started to score them. This means that some degree of reviewer bias cannot be excluded. Any reader, however, can check our points assignment and apply different weights to different criteria.

Meta-analysis is becoming fashionable. In our opinion too much stress is put on statistical pooling of study results. Criteria for deciding whether therapies, patients from different parts of the disease spectrum and end points are similar enough to be pooled are not yet available. The rationale of criteria-based meta-analyses (such as this review) is that it makes no sense to combine data from good research with data from bad research. We did not pool the results of the better studies because the material is too heterogeneous.

Double blinding, even if the placebo is described as indistinguishable, has to be checked by asking the patients in which group they believe that they were during the trial. Blindness must be checked early in the trial, before the treatment is expected to have an effect, because positive effects would break the code. It is easy to state that a trial was double blind, but patients and doctors have many ways to break the code. The Ginkgo and placebo preparations which were used in some of the trials were shown to us. They were absolutely identical in appearance, and they ought to be swallowed. However, when bitten on, the tablets can easily be distinguished, because of the bitter taste of the Ginkgo tablets, and because the contents look different. This might explain differences in favour of Ginkgo treatment. Double blinding was not checked in any trial. However, in half of all trials the placebo was described as indistinguishable (at least for several characteristics). It appears that regarding double blindness it is frequently not right to believe the authors' statements in the methodology section.

We compared the quality of the Ginkgo trials with well performed trials of pentoxifylline and co-dergocrine, which themselves are not uncriticized. Looking at the evidence of only well performed trials, it can be concluded that for intermittent claudication the state of affairs is interesting but as yet insufficient to draw positive conclusions for treatment with Ginkgo. Compared with pentoxifylline, the best trials on Ginkgo have a higher likelihood of bias. It would be interesting to see the results of new, better trials.

For cerebral insufficiency, it must be mentioned that this is a strange compilation of different symptoms. There may very well be some common factors in the etiology, but one can as easily argue that there is no connection whatsoever. We have described the characteristics of the patients in the best trials, and leave this problem to the reader. The best trials have the same quality as well performed trials of co-dergocrine.

From the mechanisms of action as described in the introduction, there seems to be some biological plausibility for potential beneficial effects of Ginkgo. Further support comes from controlled trials with healthy volunteers. Positive effects of Ginkgo were found using psychological tests for reaction and memory.^{76,77}

Another way to assess the evidence is to ask yourself whether you would take Ginkgo when you would have similar symptoms. For cerebral insufficiency, our answer would be affirmative: considering that there are no side effects, we both might try it. However, we think that also for these indications additional evidence is warranted. Vertical reading of the tables immediately shows the major drawbacks of existing research. In future experiments double blindness should be checked, and there should be larger numbers of patients, better description of randomization procedures, patient characteristics, effect measurement and data presentation.

References

1. Wagner H, Bladt S, Hartmann U, Daily A, Berkulin W. Ginkgo biloba. DC- und HPLC-Analyse von Ginkgo-Extrakten und Ginkgo-Extrakte enthaltenden Phytopräparaten. *Deutsche Apotheker Zeitung* 1989;129:2421-4.
2. Schilcher H. Ginkgo biloba L. Untersuchung zur Qualität, Wirkung, Wirksamkeit und Unbedenklichkeit. *Zeitschrift für Phytotherapie* 1988;9:119-27.
3. Drieu K. Préparation et définition de l'extrait de Ginkgo biloba. *Presse Med* 1986;15:1455-7.
4. Braquet P. Ginkgolides. Chemistry, biology, pharmacology and clinical perspectives. Barcelona: JR Prous Science Publishers, 1988.
5. Anonymous. Ginkgo biloba extract: over 5 million prescriptions a year. *Lancet* 1989;ii:1513-4.
6. Schwabe U, Paffrath D. Arzneiverordnungs-Report '89. Stuttgart: Gustav Fisher, 1989.
7. Schwabe U, Paffrath D. Arzneiverordnungs-Report '90. Stuttgart: Gustav Fisher, 1990.
8. Fünfgeld EW. Rökan Ginkgo biloba. Recent results in pharmacology and clinic. Berlin: Springer Verlag, 1988.
9. Halama P, Bartsch G, Meng G. Hirnleistungsstörungen vaskulärer Genese. Randomisierte Doppelblindstudie zur Wirksamkeit von Ginkgo-biloba-Extrakt. *Fortschr Med* 1988;106:408-13.
10. Thomson GJL, Vohra RK, Carr MH, Walker MG. A clinical trial of Ginkgo biloba extract in patients with intermittent claudication. *Int Angiol* 1990;9:75-8.
11. Kleijnen J, ter Riet G, Knipschild P. Clinical trials of homoeopathy. *Br Med J* 1991;302:316-23.
12. Bauer U. 6 Month double-blind randomised clinical trial of Ginkgo biloba extract versus placebo in two parallel groups in patients suffering from peripheral arterial insufficiency. *Arzneimittelforschung* 1984;34:716-20.
13. Bauer U. L'extrait de Ginkgo biloba dans le traitement de l'artériopathie des membres inférieurs. Etude sur 65 semaines. *Presse Med* 1986;15:1546-9.
14. Saundreau F, Serise JM, Pillet J, et al. Efficacité de l'extrait de Ginkgo biloba dans le traitement des artériopathies oblitérantes chroniques des membres inférieurs au stade III de la classification de Fontaine. *Journal des Malades Vasculaires* 1989;14:177-82.
15. Bulling B. Konservative Behandlung der chronischen peripheren arteriellen Verschlusskrankheit (PAVK) im Stadium IIb nach Fontaine mit physikalischem Training und Behandlung mit Ginkgo biloba-Extrakt EGB 761. Doppelblinde Studie im Placebo-Vergleich. Ettlingen: Intersan, 1991.
16. Salz H. Zur Wirksamkeit eines Ginkgo-biloba-Präparats bei arteriellen Durchblutungsstörungen der unteren Extremitäten. *Therapie der Gegenwart* 1980;119:1345-56.
17. Garzya G, Picari M. Trattamento delle vasculopatie periferiche con una nuova sostanza estrattiva: il Tanakan. *Clinica Europea* 1981;20(5):936-44.
18. Courbier R, Jausseran JM, Reggi M. Etude à double insu croisée du tanakan dans les artériopathies des membres inférieurs. *Méditerranée Médicale* 1977;126:61-4.
19. Böhmer D, Kalinski S, Michaelis P, Szögy A. Behandlung der PAVK mit Ginkgo-biloba-Extrakt (GBE) oder Pentoxifyllin. Eine Studie zur Beurteilung von Wirksamkeit und Verträglichkeit. *Herz Kreislauf* 1988;20:5-8.
20. Rudofsky G. Wirkung von Ginkgo-biloba-Extrakt bei arterieller Verschlusskrankheit. *Fortschr Med* 1987;105:397-400.

21. Frileux C, Copé R. L'extrait concentré de Ginkgo biloba dans les troubles vasculaires périphériques. *Cahiers d'Artériologie de Royat* 1975;3:117-22.
22. Diehm C. Placebokontrollierte multizentrische randomisiert doppel-blinde Pilotstudie zur Prüfung der Wirksamkeit des Ginkgo biloba-Extraktes EGB 761 bei Patienten mit peripherer arterieller Verschlusskrankheit im Stadium IIb nach Fontaine. Karlsruhe: Schwabe, 1990.
23. Ambrosi C, Bourde C. Nouveauté thérapeutique médicale dans les artériopathies des membres inférieurs: Tanakan. Essai clinique et étude par les cristaux liquides. *Gazette Médicale de France*. 1975;82:628-32.
24. Baitsch G. Treatment of obliterative arterial disease in diabetics (Fontaine's Stage IV) with moist gangrene. An open randomized trial of Ginkgo Biloba extract (Rökan) vs Dextran 40. *Clinical Trials Journal* 1986;23:11-9.
25. Berndt ED, Kramar M. Medikamentöse Therapie der peripheren arteriellen Verschlusskrankheit im Stadium IIb. Eine bizenrische Vergleichsstudie zwischen Ginkgo-biloba-Extrakt und Buflomedil-hydrochlorid. *Therapiewoche* 1987;37(29):2815-9.
26. Bastide G, Montsarrat M. Artérite des membres inférieurs. Intérêt du traitement médical après intervention chirurgicale. Analyse factorielle. *Gazette Médicale de France* 1978;85:4523-6.
27. Schmidt U, Rabinovici K, Lande S. Einfluß eines Ginkgo-biloba-Spezialextraktes auf die Befindlichkeit bei cerebraler Insuffizienz. *Münchener Medizinische Wochenschrift* 1991;(Suppl.):in press.
28. Brüchert E, Heinrich SE, Ruf-Kohler P. Multizentrische Doppelblindstudie des Fachverbandes Deutscher Allgemeinärzte zur Prüfung der Wirksamkeit von LI 1370 bei älteren Patienten mit Hirnleistungsschwäche. *Münchener Medizinische Wochenschrift* 1991;(Suppl.):in press.
29. 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.
30. Taillandier J, Ammar A, Rabourdin JP, et al. Traitement der troubles du vieillissement cérébral par l'extrait de Ginkgo biloba. *Presse Med* 1986;15:1583-7.
31. 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.
32. Vorberg G, Schenk N, Schmidt U. Wirksamkeit eines neuen Ginkgo-biloba-Extraktes bei 100 Patienten mit zerebraler Insuffizienz. *Herz + Gefäße* 1989;9(7):936-41.
33. Eckmann F. Hirnleistungsstörungen - Behandlung mit Ginkgo-biloba-Extrakt. *Fortschr Med* 1990;108:557-60.
34. 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.
35. Augustin P. Le Tanakan en gériatrie. Etude clinique et psychométrique chez 189 malades d'hospice. *Psychologie Médicale* 1976;8:123-30.
36. Arrigo A. Behandlung der chronischen zerebrovaskulären Insuffizienz mit Ginkgo-biloba-Extrakt. *Therapiewoche* 1986;36:5208-18.
37. Arrigo A, Cattaneo S. Clinical and psychometric evaluation of Ginkgo biloba extract in chronic cerebro-vascular diseases. In: Agnoli, Rapin JR, Scapagnini V, Weitbrecht WV. Effects of Ginkgo Biloba Extract on Organic Cerebral Impairment. John Libbey Eurotext Ltd, 1985: 85-9.
38. Halama P. Ginkgo biloba. Wirksamkeit eines Spezialextraktes bei Patienten mit zerebraler Insuffizienz. *Münchener Medizinische Wochenschrift*. 1991;133:190-4.
39. Meyer B. Étude multicentrique des acouphènes. *Epidémiologie et thérapeutique. Ann Otolaryngol Chir Cervicofac* 1986;103:185-8.
40. Chesseboeuf L, Herard J, Trevin J. Étude comparative de deux vasorégulateurs dans les hypoacusies et les syndromes vertigineux. *Médecine du Nord et de l'Est* 1979;3(5):534-9.
41. Maier-Hauff K. Wirksamkeit von LI 1370 bei ambulanten Patienten nach cerebraler Aneurysma-Operation mit Störungen der Hirnleistungsfähigkeit. *Münchener Medizinische Wochenschrift* 1991;(Suppl.):in press.
42. Weitbrecht W-U, Jansen W. Primär degenerative Demenz: Therapie mit Ginkgo-biloba-Extrakt. *Fortschr Med* 1986;104:199-202.
43. Weitbrecht W-U, Jansen W. Doubleblind and comparative (Ginkgo biloba versus placebo) therapeutic study in geriatric patients with primary degenerative dementia - a preliminary evaluation. In: Agnoli, Rapin JR, Scapagnini V, Weitbrecht WV. Effects of Ginkgo Biloba Extract on Organic Cerebral Impairment. John Libbey Eurotext Ltd, 1985:91-9.
44. Schwerdtfeger F. Elektronystagmographisch und klinisch dokumentierte Therapieerfahrungen mit rökan bei Schwindelsymptomatik. *Therapiewoche* 1981;31:8658-67.
45. Moreau P. Un nouveau stimulant circulatoire cérébral. *Presse Med* 1975;4:2401-2.
46. Dieli G, La Mantia V, Saetta M, Costanzo E. Essai clinique à double insu du Tanakan dans l'insuffisance cérébrale chronique. *Lavoro Neuropsichiatrico* 1981;68:1-10.
47. Eckmann F, Schlag H. Kontrollierte Doppelblind-Studie zum Wirksamkeitsnachweis von Tebonin forte bei Patienten mit zerebrovaskulärer Insuffizienz. *Fortschr Med* 1982;100:1474-8.
48. 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.

49. Hofferberth B. Prüfung der Wirksamkeit von Ginkgo-biloba-Spezialextrakt mit neurophysiologischen und psychometrischen Methoden bei Patienten mit hirnorganischem Psychosyndrom. Münchener Medizinische Wochenschrift 1991;(Suppl.):in press.
50. Hofferberth B. Simultanerfassung elektrophysiologischer, psychometrischer und rheologischer Parameter bei Patienten mit hirnorganischem Psychosyndrom und erhöhtem Gefäßrisiko. Eine Placebo-kontrollierte Doppelblindstudie mit Ginkgo biloba-Extrakt EGB 761. In: Stodtmeister R, Pillunat LE (Eds). Mikrozirkulation in Gehirn und Sinnesorganen. Stuttgart: Ferdinand Enke Verlag, 1991:64-74.
51. Hofferberth B. Einfluß von Ginkgo biloba-Extrakt auf neurophysiologische Meßergebnisse bei Patienten mit hirnorganischen Psychosyndrom. Eine Doppelblindstudie gegen Placebo. Arzneimittelforschung 1989;39:918-22.
52. Israël L, Dell'Accio E, Martin G, Hugonot R. Extrait de Ginkgo biloba et exercices d'entraînement de la mémoire. Évaluation comparative chez des personnes âgées ambulatoires. Psychologie Médicale 1987;19(8):1431-9.
53. Haan J, Reckermann U, Welter FL, Sabin G, Müller E. Ginkgo-biloba-Flavonglykoside. Therapiemöglichkeit der zerebralen Insuffizienz. Medizinische Welt 1982;33:1001-5.
54. Claussen CF. Mit Ginkgo biloba wird Ihr Patient wieder standfest. Ärztliche Praxis 1984;36:193-4.
55. Gerhardt G, Rogalla K, Jaeger J. Medikamentöse Therapie von Hirnleistungsstörungen. Randomisierte Vergleichstudie mit Dihydroergotoxin und Ginkgo-biloba-Extrakt. Fortschr Med 1990;108:384-8.
56. Hartmann A, Frick M. Wirkung eines Ginkgo-biloba-Spezialextraktes auf psychometrische Parameter bei Patienten mit vaskulär bedingter Demenz. Münchener Medizinische Wochenschrift 1991;(Suppl.):in press.
57. 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.
58. Agnoli A. Relazione clinica sulla specialita' Tebonin forte. Milano: ALSO Lab. S.a.S., 1980.
59. Hamann KF. Physikalische Therapie des vestibulären Schwindels in Verbindung mit Ginkgo-biloba-Extrakt. Eine posturografische Studie. Therapiewoche 1985;35:4586-90.
60. Teigeler R, Pieprzyk L. Ginkgo biloba bei zerebraler Insuffizienz. Ärztliche Praxis 1984;36:1374.
61. 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.
62. Chartres JP, 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.
63. Natali R, Rachinel J, Pouyat PM. Essai comparatif croisé en O.R.L. de deux médications vaso-actives. Cahiers d'Oto Rhino Laryngologie 1979;14(2):185-90.
64. Bono Y, Mouren P. L'insuffisance circulatoire cérébrale et son traitement par l'extrait de Ginkgo biloba. Méditerranée Médicale 1975;3(61):59-62.
65. 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 Médical 1977;13:1197-9.
66. Lindgärde F, Jelnes R, Björkman H, Et al. Conservative drug treatment in patients with moderately severe chronic occlusive peripheral arterial disease. Circulation 1989;80:1549-56.
67. Tonak J, Knecht H, Groitt H. Zur Behandlung von Durchblutungsstörungen mit Pentoxifyllin. Eine Doppelblindstudie mit Trental. Medizinische Monatsschrift 1977;31:467-72.
68. Porter JM, Cutler BS, Lee BY, et al. Pentoxifylline efficacy in the treatment of intermittent claudication: Multicenter controlled double-blind trial with objective assessment of chronic occlusive arterial disease patients. Am Heart J 1982;104:66-72.
69. Kellner H. Treatment of chronic arterial circulatory disorders: a double-blind trial with pentoxifylline (Trental 400). Pharmatherapeutica 1983;3(Suppl.1):67-73.
70. Roekaerts F, Deleers L. Trental 400 in the treatment of intermittent claudication: results of long-term, placebo-controlled administration. Angiology 1984;35:396-406.
71. Lazzari R, Passeri M, Chierichetti SM. Le mésylate de dihydroergotoxine dans le traitement de l'insuffisance cérébrale sénile. Résultat d'une étude clinique multicentrique en double aveugle contre placebo à long terme. Presse Med 1983;12:3179-85.
72. Yoshikawa M, Hirai S, Aizawa T, et al. A dose-response study with dihydroergotoxine mesylate in cerebrovascular disturbances. J Am Geriatr Soc 1983;31:1-7.
73. Bargheon J. Étude en double insu de l'hydergine chez le sujet âgé. Presse Med 1973;2:2053-5.
74. McConnachie RW. A clinical trial comparing 'Hydergine' with placebo in the treatment of cerebrovascular insufficiency in elderly patients. Curr Med Res Opinion 1973;1(8):463-8.
75. Hollingsworth SW. Response of geriatric patients from the satellite nursing homes of Maricopa county to Hydergine therapy: A double-blind study. Curr Ther Res 1980;27:401-10.
76. Schaffler K, Reeh PW. Doppelblindstudie zur hypoxieprotektiven Wirkung eines standardisierten Ginkgo-Biloba-Präparates nach Mehrfachverabreichung an gesunden Probanden. Arzneimittelforschung 1985;35:1283-6.
77. Subhan Z, Hindmarch I. The psychopharmacological effects of Ginkgo biloba extract in normal healthy volunteers. Int J Clin Pharmacol Res 1984;4:89-93.

CHAPTER 12

CLINICAL TRIALS OF HOMOEOPATHY

Clinical trials of homoeopathy

Jos Kleijnen, Paul Knipschild, Gerben ter Riet

Abstract

Objective—To establish whether there is evidence of the efficacy of homoeopathy from controlled trials in humans.

Design—Criteria based meta-analysis. Assessment of the methodological quality of 107 controlled trials in 96 published reports found after an extensive search. Trials were scored using a list of predefined criteria of good methodology, and the outcome of the trials was interpreted in relation to their quality.

Setting—Controlled trials published world wide.

Main outcome measures—Results of the trials with the best methodological quality. Trials of classical homoeopathy and several modern varieties were considered separately.

Results—In 14 trials some form of classical homoeopathy was tested and in 58 trials the same single homoeopathic treatment was given to patients with comparable conventional diagnoses. Combinations of several homoeopathic treatments were tested in 26 trials; isopathy was tested in nine trials. Most trials seemed to be of very low quality, but there were many exceptions. The results showed a positive trend regardless of the quality of the trial or the variety of homoeopathy used. Overall, of the 105 trials with interpretable results, 81 trials indicated positive results whereas in 24 trials no positive effects of homoeopathy were found. The results of the review may be complicated by publication bias, especially in such a controversial subject as homoeopathy.

Conclusions—At the moment the evidence of clinical trials is positive but not sufficient to draw definitive conclusions because most trials are of low methodological quality and because of the unknown role of publication bias. This indicates that there is a legitimate case for further evaluation of homoeopathy, but only by means of well performed trials.

Introduction

A survey of 293 general practitioners in The Netherlands showed that 45% of them think that homoeopathic remedies are efficacious in treating upper respiratory tract infections or hay fever.¹ On the other hand, many doctors do not believe that homoeopathy is an efficacious treatment as it is highly implausible that infinitesimally diluted substances retain their biological effects. It is also often stated that homoeopathy has not been evaluated using modern methods—that is, controlled trials. The first argument may be true, but the second is certainly not true. Reading an article about pollen C30 in hay fever increased our interest in homoeopathy.² We could not believe the positive result (was it coincidence?) and therefore we started to search for further reports. Here we present 107 controlled trials of homoeopathy.

Homoeopathic medicine is a system developed by Samuel Hahnemann from the *similia* concept: "*similia similibus curantur*." This implies that a diluted, "potentised" agent, which (when undiluted) in healthy

Department of
Epidemiology and Health
Care Research, University
of Limburg, PO Box 616,
3000 MD Maastricht, The
Netherlands

Jos Kleijnen, MD, research
fellow
Paul Knipschild, MD,
professor of epidemiology
Gerben ter Riet, MD, research
fellow

Correspondence to:
Dr Kleijnen.

107 1991;302:316-23

individuals induces complaints resembling those of the patient, can be used to cure the patient.

Potentiation is a combination of dilution and shaking of a substance. A plant—for example, *Arnica montana*—is macerated and dissolved in alcohol. One part of this “mother tincture” is mixed with nine parts (D1 potency) or 99 parts (C1 potency) of 90% alcohol (the concentration of the alcoholic solution may vary between manufacturers) and then vigorously shaken. This process can be repeated many times, resulting in very high dilutions (potencies): D6 means one molecule of the original substance in 10^6 molecules of 90% alcohol; C6 means one molecule in 10^{12} molecules. In potencies of D24 or C12 and higher it is very unlikely that even a single molecule of the mother tincture is present. The idea is, however, that higher potencies work more strongly than lower potencies.

Using the *similia* principle the classical homeopathy tries to find a substance that fits the patient's complaints as much as possible. Unusual symptoms that do not fit the symptom complexes recognised by conventional medicine may be considered even more important than the regular symptoms. This is why homeopathy is a highly individualised treatment, resulting in different treatments for patients who would receive an identical treatment in conventional medicine. In modern homeopathy combinations of several or many homeopathic substances are often used, especially in over the counter preparations. The classical homeopathy will never use this polypharmacy. Also, according to classical homeopathy a *similium* must be used and not a potentiation of the causal agent (for example, pollen in hay fever or lead in lead poisoning), which is called *isopathy*. Phytotherapy is the administration of herbs or low potencies of herbs (D2 or so). These preparations may still have pharmacological effects, and therefore it is sometimes difficult to demarcate phytotherapy from modern homeopathy, the fundamental difference being the applied low dose toxicology principle in homeopathy. This description of homeopathy indicates that it is not just another therapy but a distinct outlook in medicine, and several interpretations have developed, often contradictory to one another.

For this review we searched exhaustively for published reports to investigate the clinical evidence of the efficacy of homeopathy, regardless of its (to us) implausibility. The positive and negative evidence was weighed against the methodological quality of the research.

Materials and methods

Trials were eligible if parallel index and control groups were included. Crossover designs were also eligible, but controlled studies in animal models were excluded.

Experiments were found by various strategies: a computer search (MEDLINE online 1966-90; keyword homeopathy); checking references extensively, in articles on clinical research and in textbooks¹⁻³; checking the proceedings of conferences of homeopathy; checking the contents of several journals of homeopathy; personal communication with researchers; writing to and visiting major manufacturers of homeopathic preparations; and visiting several libraries specialising in homeopathy. This process of collection took place over a period of more than three years. Trials published in any language were eligible, without restrictions.

Classical homeopathy uses individual diagnoses and treatments. From a homogeneous group given diagnoses in conventional medicine the patients suitable for homeopathic treatment can be selected. This results in acceptable participants from both regular

and homeopathic points of view. Individual treatment is prescribed, and then the patients are randomly allocated to homeopathic or placebo treatment. If necessary, the prescription may be changed in the course of time and, of course, patients who started on placebo stay on placebo.⁴

When the same homeopathic drug or combination of homeopathic drugs is given to all patients with a comparable regular diagnosis, trial methodology is the same as in regular medicine. This also goes for trials testing isopathy.

Because the effects of most homeopathic treatments are meant to last for longer periods, the interpretation of crossover trials is complicated by carryover effects. The analysis will be very difficult, and consequently parallel experiments are preferable.

To explore the possibility that an increasing likelihood of bias (an increasing number of methodological shortcomings) is reflected in the results of the trials, criteria for a methodological assessment of the experiments were established. We put much weight on the number of participants. In most indications for homeopathic treatment subjective symptoms are the main outcome phenomenon. Substantial improvements of patients in the control group can be expected, and fairly large groups, which are comparable at baseline for prognostic factors, are needed for valid assessment of the efficacy. In trials with limited numbers of participants one cannot be confident that randomisation will equally divide known and unknown confounders over the experimental and control groups. As well, publication bias may be less likely for experiments with large numbers of participants: the effort and costs entailed will increase the likelihood that a paper is submitted for publication. Thus a main argument for our emphasis on relatively large numbers of participants was not the likelihood of type II error, which also depends on the estimated size of the effect, but mainly our worry about incomparability at baseline of the groups and the likelihood of publication bias.

Other major criteria for methodological soundness were randomisation and double blindness. When prognostic factors of the illness, other than the intervention under study, are insufficiently known, random allocation to the contrasted treatments is useful to ensure a comparable prognosis. Double blindness is important for keeping the intervention exactly the same in the contrasted groups except for the homeopathic treatment, and for an unbiased assessment of the effects. This is especially important if it concerns the relief of subjective symptoms, as is often the case in homeopathic treatment.

Starting from a maximum score of 100 points, we divided these among seven methodological criteria.

(1) *Patient characteristics adequately described: 10 points*—Description of the symptoms and, if appropriate, of their duration and severity.

(2) *Number of patients analysed: 30 points*—One hundred or more patients per group analysed=30 points, 50-99 patients per group=20 points, and 25-49 patients per group=10 points. A crossover trial with 70 participants (35 given active treatment and 35 given placebo in each period) would score 10 points. In trials assessing the prophylactic effects of homeopathy the number of patients with the outcome phenomenon was used.

(3) *Randomisation: 20 points*—Twenty points if the method of randomisation was described and correct, 10 points if the method was not described or if some form of pseudorandomisation was applied. If there were fewer than 25 participants per group, half the score was given unless there was prestratification (matching) on relevant items and a table showing comparable baseline characteristics.

(4) *Intervention well described: 5 points*—Adminis-

tration (doses, duration) and origin (method of manufacture) of homeopathic preparations.

(5) *Double blinding: 20 points*—Twenty points if the placebo was described as indistinguishable, 10 points if double blinding was only mentioned.

(6) *Effect measurement relevant and well described: 10 points*—Measurement of the effect must be sensible and reproducible. Five points each for relevance and adequate description.

(7) *Presentation of the results in such a manner that the analysis can be checked by the reader: 5 points*—Depending on measurement of the effect, at least the mean(s) and standard deviation, standard error, or confidence interval in each group must be mentioned, or the number of patients with a certain outcome (for example, if rates or proportions were used).

Sometimes only part of the score was given if the description was unclear, or if only some of several interventions, measurements of outcome, or data presentations met the criteria. In the second criterion we chose to use the number of patients analysed instead of the number randomised because in many publications drop outs were not accounted for. Often the number of patients admitted was not even mentioned. In the seventh criterion we did not demand confidence intervals for the comparisons between groups because then virtually no trials would score the criterion, with only a few exceptions.^{1,19}

All articles were scored by at least two of us, and differences, which were mainly caused by reading errors or by unclear descriptions in the publications, were resolved by discussions. Most of these differences occurred in patient characteristics and descriptions of measurement of the effect; in these cases the relevance and sensibility had to be judged. The largest difference was 13 points.

Assessment of articles using these criteria provides a score that gives an indication of the methodological quality of each trial. This quality is an important factor in weighing the conclusions of different trials and, of course, on the impact on the reader's opinion of all the evidence presented. We have selected well established methodological criteria,¹⁰ and our assessment can be checked by the reader (table 1).

Results

Table 1 shows some methodological characteristics of the better trials (those scoring 55 points or more).^{1,13-31} Some good studies have been reported, but overall the methodological quality was disappoint-

ing. Patient characteristics were described adequately in 56 trials. More than half of the publications (63) were of trials in which fewer than 25 patients per group were treated. Sixty eight trials were randomised, but only 17 described the method of randomisation. The intervention was adequately or reasonably well described in 80 trials. Seventy five were double blind, but the placebo was described as indistinguishable in only 31 trials. In 67 publications the effect measurement was judged to have been sensible and well described. Sufficient data for the reader to check the analysis were given in 65 trials.

It is difficult to compare the quality of trials that score more or less the same, and in the lower range the results of all studies may be seriously biased because of several methodological shortcomings. Consequently, we present in detail the results of only the best trials (those scoring 60 points or more) (table 1).¹¹⁻¹⁴

In 14 experiments some form of classical homeopathy was tested.²²⁻²⁴ Only one of these scored more than 60 points. In a randomised double blind trial Brigo gave one or sometimes two of eight chosen drugs (belladonna, gelsemium, ignatia, cyclamen, lachesis, natrium muriaticum, silica, or sulphur in a C30 potency) to 30 patients with migraine headache; 30 controls received a placebo. After four months the patients treated with homeopathy fared much better than the controls on severity of attacks: on a 10 cm visual analogue scale the severity changed from 9.1 to 2.9 in the homeopathic group and from 8.4 to 7.8 in the control group. Similar differences were found for the frequency and the duration of the attacks.²²

In about half of the controlled trials (58 studies) the same single homeopathic treatment was given to a group of patients with comparable conventional diagnoses. Combinations of homeopathic treatments (polypharmacy) were tested in 26 studies, and isopathy in nine. Only one trial compared dilutions with potencies (a positive trend was found in favour of the potency)¹¹ and in a few trials different potencies or different homeopathic substances were compared with each other.^{12,15,21,26-29}

Twenty eight trials were published before 1980, 38 in the period 1980-4 and 41 from 1985 onwards. Forty two trials were published in English, 34 in German, 30 in French, one in Italian, and one in Portuguese. Several trials were published in more than one language (for example, Italian and French); in those cases we chose the reference of the most comprehensive and most easily obtainable publication.

According to conventional diagnoses, several groups

TABLE 1—Scoring of methodological characteristics of clinical trials of homeopathy

	Characteristics of patients (max=10)	Number of patients analysed (max=30)	Randomisation (max=20)	Intervention (max=5)	Double blinding (max=20)	Measurement of effect (max=10)	Presentation of data (max=5)	Total score (max=100)
GRECHO 1989 ¹¹	10	30	10	5	20	10	5	90
Reilly <i>et al</i> 1986 ¹²	10	20	20	5	20	10	5	90
Ferley <i>et al</i> 1989 ¹³	10	30	10	5	20	8	5	88
Wiesnauer <i>et al</i> 1985 ¹⁴	5	20	20	5	20	10	5	85
Arnal-Lasserre 1986 ¹⁵	10	10	20	5	20	10	5	80
Wiesnauer and Gaus 1986 ¹⁶	10	20	10	5	20	10	5	80
Zell <i>et al</i> 1988 ¹⁷	10	20	20	5	10	10	5	80
Valero (Raphanus sativus) 1981 ¹⁸	10	30	10	5	10	10	0	75
Aulagnier 1985 ¹⁹	5	10	20	5	20	10	5	75
Wiesnauer <i>et al</i> 1983 ²⁰	10	10	10	5	20	10	5	70
Bordes and Durtman 1986 ²¹	10	10	20	5	10	10	5	70
Valero (Pyrogenum) 1981 ²²	8	10	10	5	10	10	5	68
Ferley <i>et al</i> 1987 ²³	10	10	20	3	10	10	5	68
Brigo 1987 ²⁴	10	20	15	5	0	10	5	65
Maiwald <i>et al</i> 1988 ²⁵	10	10	10	5	20	10	0	60
Wiesnauer <i>et al</i> 1989 ²⁶	5	10	10	3	20	10	5	58
Bignamini <i>et al</i> 1987 ²⁷	10	10	10	3	10	10	5	58
Chevrel <i>et al</i> 1984 ²⁸	10	10	20	3	0	10	5	58
Gassinger <i>et al</i> 1981 ²⁹	5	20	10	3	10	5	5	58
Ritter 1966 ³⁰	10	0	10	5	20	10	5	55
Wiesnauer and Gaus 1987 ³¹	10	0	5	5	20	10	5	55
Savage 1977 ³²	10	0	5	5	20	10	5	55

Eighty four controlled trials scored <55 points.^{1,11-10}

TABLE II—Characteristics and results of best trials

	Score for methodology (max = 100)	Indication (No of patients/No of controls)	Intervention	Results (No of patients/No of controls)
Polypharmacy:				
Ferley <i>et al</i> 1989 ^a	88	Treatment of influenza (237/241)	Anas barbariae hepatitis, cordis extractum C200 v placebo	Recovery rate within 48 hours (17.1%/10.3%)
Arnal-Lasserre 1986 ^a	80	Duration of delivery (53/40)	Actea racemosa C5, arnica C5, caulophyllum C5, gelsemium C5, pusailla C5 v placebo	Duration of delivery: (5.1/8.5 hours); "dystocie" [problems with dilatation] (11.3%/40%)
Zell <i>et al</i> 1988 ^a	80	Ankle sprains (33/36)	D2-D6 combination of 14 substances v placebo	No of patients without pain after 10 days: (28/13)
Aulagnier 1985 ^a	75	Bowel movements after abdominal operation (100/100)	Opium C9, raphanus C9, arnica C9 v placebo	Days until first flatus (2.5/3.2); days until first faeces (4.0/4.9)
Bordes and Dorfman 1986 ^a	79	Dry cough (30/30)	C3 combination of 10 substances v placebo	Very good or good result after 1 week (20/8)
Ferley <i>et al</i> 1987 ^a	68	Prevention and treatment of influenza (588/594)	D1-D6 combination of 10 substances v placebo	Incidence (6.5%/7.2%); duration of symptoms (7.0/6.8 days)
Maiwald <i>et al</i> 1988 ^a	65	Influenza (88/82)	Aconitum D4, bryonia D4, lachesis D12, eupatorium perfoliatum D3, phosphorus D5 v acetyl salicylic acid 1500 mg days 1-4, 500 mg days 5-10	Positive result within 4 days (29%/23%)
Wiesnauer <i>et al</i> 1989 ^a	60	Sinusitis (45, 38, 35/34)	(1) Luffa operculata D4, kalium bichromicum D4, cinnabaris D3 (2) Kalium bichromicum D4, cinnabaris D3 (3) Luffa operculata D4; v (4) placebo	Combination score of 6 symptoms (no difference between the 4 groups)
Same formula in all patients:				
GRECHO 1989 ^a ^b	90	Bowel movements after abdominal operation (4 groups of 150)	(1) Opium C15 (2) Opium C15, raphanus C5 v (3) Placebo (4) No treatment	Time until first faeces: (1) 96 hours (2) 99 hours (3) 94 hours (4) 95 hours Similar results for first peristaltic sounds and first flatus
Wiesnauer and Gaus 1985 ^a	85	Pollinosis (50/55, 59)	(1) Galphimia glauca D6 v (2) Galphimia glauca dilution 10 ⁻⁴ (3) Placebo	Improvement of nasal symptoms after 2, 4 weeks: (1) 60%, 78% (2) 40%, 51% (3) 41%, 58% Similar results for ocular symptoms
Valero 1981 ^a	80	Postoperative infections (54/74)	Raphanus C7 v placebo	No of patients with infection (15/20)
Valero 1981 ^a	70	Bowel movements after abdominal operation (43/37)	Pyrogenium C7 v placebo	Time until first flatus (53.3/58.6 hours)
Wiesnauer <i>et al</i> 1983 ^a	75	Pollinosis (41/45)	(1) Galphimia glauca D4 v (2) Placebo	Improvement of symptoms after 2, 4 weeks: (1) 83%, 81% (2) 47%, 57%
Comparison of several homeopathic treatments:				
Wiesnauer and Gaus 1986 ^a	80	Pollinosis (62, 56, 54, 63)	Galphimia glauca (1) C2 (2) C4 (3) D4 (4) LM4	Improvement of nasal symptoms after 2, 4 weeks: (1) 67%, 83% (2) 71%, 79% (3) 67%, 82% (4) 69%, 85% Improvement of ocular symptoms after 2, 4 weeks: (1) 64%, 83% (2) 73%, 88% (3) 65%, 82% (4) 76%, 89%
Isopathy:				
Reilly <i>et al</i> 1986 ^a	90	Pollinosis (74/70)	Pollen C30 v placebo	Change in 100 mm visual analogue scale symptom score after 5 weeks (-17.2 mm/-2.6 mm)
Classical homeopathy:				
Brigo 1987 ^a	68	Migraine (30/30)	8 possible homeopathic remedies C30 v placebo	Change in 10 cm visual analogue scale symptom score after 4 months (-8.2 cm/-0.6 cm). Similar results for frequency and duration of attacks

of indications emerged: diseases of the respiratory system (19 trials on respiratory infections, five trials on hay fever, and one on asthma); gastrointestinal complaints (seven trials); and pain from several sources (27 trials, of which six were of rheumatological diseases). Table III presents the outcome of all 107 trials. In 42 we thought that insufficient data were given to check the authors' interpretation of the outcome(s). Consequently the results reflect not our conclusions but the inference made by the authors of the publications, who to us seem sometimes to be a little overoptimistic. In most cases, however, a positive result indicates that there was a statistically significant difference in the main outcome(s) between the contrasted groups, whereas a negative result means that no significant difference was found ($p > 0.05$). We could not pool the results statistically because of the heterogeneity of the studies.

The evidence is to a large extent positive: of the better studies 15 trials showed positive results whereas in seven trials no positive effect could be detected (in one trial only homeopathic treatments were compared with each other). The trials with a methodological score below 55 points showed an even clearer trend: in most publications positive results were reported (66 positive, 17 negative). Overall, of the 105 trials with interpretable results, 81 indicated positive results whereas in 24 trials no positive effects of homeopathy were found compared with (mostly) placebo controls. In the two other trials only homeopathic treatments were compared to each other.

Discussion

In the methods section we indicated that it is possible to perform trials on the efficacy of homeo-

pathy, including classical homoeopathy, in a way that is acceptable for both sceptical physicians and enthusiastic homoeopaths. Criticisms of these methods, often suggesting that special methodology and statistics are needed for the evaluation of homoeopathy, are in our opinion based on lack of knowledge of research methodology.

A problem in our methodological assessment is that limited description of the methods and the results in the publication may lead to a lower score. We believe, however, that a detailed description of this information is as important as using good methodology in practice. It could be argued that other criteria should be used for the methodological assessment and that this kind of assessment is rather subjective. As stated before, we

have selected well established criteria. The reader could apply different weights to the criteria to see whether substantial changes would occur in our methodological ranking, but we think that this will not be the case.

Double blinding, even if the placebo is described as indistinguishable, has to be checked by asking the patients in which group they believe that they were during the trial. Blindness must be checked early in the trial, before the treatment is expected to take effect, because positive effects would break the code. It is easy to state that a trial was double blind, but patients have many ways to break the code. This might explain small differences in favour of homoeopathy. Double blinding was not checked in any trial of homoeopathy.

TABLE III—Clinical trials of homoeopathy grouped according to diagnoses from conventional medicine

Indication	Score (max = 100)	Result	Indication	Score (max = 100)	Result
Diseases of the vascular system:			Rheumatological disease:		
Bignamini <i>et al</i> 1987 ²³	Hypertension	58 Negative	Shipley <i>et al</i> 1983 ³	Osteoarthritis	50 Negative
Wiesenaue and Gaus 1987 ²⁴	Hypotension	58 Positive	Fisher <i>et al</i> 1989 ³	Fibromyalgia	45 Positive
Savage 1977 ²⁵	Stroke	55 Negative	Gibson <i>et al</i> 1980 ³	Rheumatoid arthritis	40 Positive
Gauthier 1983 ²⁶	Flushing	53 Negative	Audrade <i>et al</i> 1988 ³	Rheumatoid arthritis	38 Negative
Savage and Roe 1978 ²⁷	Stroke	53 Negative	Fisher 1986 ³	Fibrositis	38 Positive
Hitzzenberger <i>et al</i> 1982 ²⁸	Hypertension	48 Negative	Gibson <i>et al</i> 1978 ³	Rheumatoid arthritis	33 Positive
Dorfman <i>et al</i> 1988 ²⁹	Venous perfusion	35 Positive			
Hadjicostas <i>et al</i> 1988 ³⁰	Bleeding	35 Positive	Trauma or pain:		
Master 1987 ³²	Hypertension	13 Positive	Zell <i>et al</i> 1988 ³¹	Ankle sprains	80 Positive
			Brigo 1987 ³²	Migraine	68 Positive
Respiratory infections:			Bourgeois 1984 ³³	Haematoma	53 Positive
Ferley <i>et al</i> 1989 ³⁵	Influenza	38 Positive	Casanova 1981 ³⁴	Myalgia	45 Positive
Bordes and Dorfman 1986 ³⁶	Coughing	70 Positive	Pinsent <i>et al</i> 1986 ³⁵	Dental extraction	45 Positive
Ferley <i>et al</i> 1987 ³¹	Influenza	68 Negative	Berthier 1985 ³⁶	Dental extraction	40 Positive
Maiwald <i>et al</i> 1988 ³⁸	Influenza	65 Positive	Albertini <i>et al</i> 1984 ³⁷	Dental neuralgia	38 Positive
Wiesenaue <i>et al</i> 1989 ³⁹	Sinusitis	60 Negative	Campbell 1976 ³⁸	Bruising	38 Negative
Gastinger <i>et al</i> 1981 ⁴⁰	Common cold	58 Positive	Hildebrand and Eltze 1983 ³⁹	Myalgia	38 Positive
Lewith <i>et al</i> 1989 ⁴¹	Influenza	55 Negative	Hildebrand and Eltze 1983 ³⁹	Myalgia	38 Positive
Lecocq 1985 ⁴²	Respiratory infections	50 Positive	Hildebrand and Eltze 1983 ³⁹	Myalgia	38 Positive
Lewis 1984 ⁴³	Whooping cough	49 Negative	Hildebrand and Eltze 1983 ³⁹	Myalgia	38 Positive
Schmidt 1987 ⁴⁴	Bronchitis	45 Positive	Leaman and Gorman 1989 ⁵⁰	Minor burns	38 Negative
Chakravarty <i>et al</i> 1977 ⁴⁵	Tonsillitis	38 Positive	Geiger 1968 ⁵¹	Oedema	35 Positive
Mossinger 1985 ⁴⁶	Otitis media	38 Positive	Kubista <i>et al</i> 1986 ⁵²	Mastalgia	35 Positive
Davies 1977 ⁴⁷	Influenza	35 Positive	Michaud 1981 ⁵³	Oedema	35 Positive
Mossinger 1973 ⁴⁸	Pharyngitis	35 Positive	Mergen 1969 ⁵⁴	Oedema	33*
Mossinger 1982 ⁴⁹	Common cold	35 Negative	Caspar and Foerstel 1967 ⁵⁵	Oedema	28 Positive
Hourest 1982 ⁵¹	Respiratory infections	28 Positive	Campbell 1976 ⁵⁶	Bruising	28 Positive
Mossinger 1976 ⁵²	Pharyngitis	25 Positive	Khan 1985 ⁵⁷	Haflax valgus	15 Positive
Alasciello and Felesi 1985 ⁵³	Influenza	18 Positive	Anonymous 1980 ⁵⁸	Cystitis	13 Positive
Bungetzianu 1988 ⁵⁴	Influenza	0 Negative			
Other infections:			Mental or psychological problems:		
Valero 1981 ⁵⁷	Postoperative infection	80 Negative	Delaunay 1983 ⁶¹	Behaviour in children	48 Positive
Valero 1981 ⁵⁸	Postoperative infection	50 Positive	Carlini <i>et al</i> 1987 ⁶²	Insomnia	45 Negative
Ustianowski 1974 ⁵⁹	Cystitis	45 Positive	Heulluy 1985 ⁶³	Depression	45 Positive
Mossinger 1980 ⁶⁰	Furuncles	43 Positive	Ponti 1986 ⁶⁴	Travel sickness	40 Positive
Subramanyam <i>et al</i> 1990 ⁶¹	Filaria	38 Positive	Tsiakopoulos <i>et al</i> 1988 ⁶⁵	Vertigo	35 Positive
Carey 1986 ⁶²	Vaginal discharge	35 Positive	Vu Din Sao and DeLaunay 1983 ⁶⁶	Nervous tension	30 Positive
Castro and Nogueira 1975 ⁶⁴	Meningitis	13 Positive	Dexpert 87 ⁶⁷	Seasickness	25 Positive
Diseases of the digestive system:			Alibeu and Jobert 1990 ⁷⁰	Agitation	23 Positive
Ritter 1966 ⁶⁶	Gastritis	58 Positive	Davies 1988 ⁷¹	Aluminium deficiency	23 Negative
Rahlfis and Mossinger 1979 ⁶⁸	Irritable colon	50 Positive	Master 1987 ⁷²	Aphasia	23 Positive
Owen 1990 ⁶⁹	Irritable colon	35 Positive			
Rahlfis and Mossinger 1976 ⁷²	Irritable colon	35 Positive	Other diagnoses:		
Mossinger 1976 ⁷³	Abdominal complaints	23 Negative	Arnal-Laserre 1986 ⁷⁶	Duration of delivery	80 Positive
Mossinger 1974 ⁷⁴	Cholecystopathy	15 Positive	Skaliotis <i>et al</i> 1988 ⁷⁷	Diabetes	50 Positive
Mossinger 1976 ⁷⁵	Abdominal complaints	13 Negative	Coudert-Deguilhaume 1981 ⁷⁸	Duration of delivery	45 Positive
Pollinosis:			Kennedy 1971 ⁷⁹	Postoperative complications	45 Negative
Reilly <i>et al</i> 1986 ⁷⁸	Pollinosis	90 Positive	Pateron 1943 ⁸⁰	Gas poisoning	41 Positive
Wiesenaue and Gaus 1985 ⁷⁹	Pollinosis	85 Positive	Basu 1980 ⁸¹	Myopia	35 Positive
Wiesenaue and Gaus 1986 ⁸⁰	Pollinosis	80 *	Harivau 1987 ⁸²	Cramps/dysuria	35 Positive
Wiesenaue <i>et al</i> 1983 ⁸¹	Pollinosis	75 Positive	Kirchhoff 1982 ⁸³	Lymphoedema	33 Positive
Reilly and Taylor 1985 ⁸²	Pollinosis	50 Positive	Kienle 1973 ⁸⁴	Respiratory insufficiency	30 Positive
Reilly <i>et al</i> 1990 ⁸³	Asthma	35 Positive	Pateron 1943 ⁸⁵	Gas poisoning	28 Positive
Recovery of bowel movements after surgery:			Ventokovskoy and Popov 1990 ⁸⁶	Complications of delivery	22 Positive
GRECHO 1989 ⁸⁶	Ileus	90 Negative	Schwab 1990 ⁹⁰	Skin diseases	20 Positive
Aufagnier 1985 ⁸⁷	Ileus	75 Positive	Schwab 1990 ⁹¹	Skin diseases	20 Positive
Valero 1981 ⁸⁸	Ileus	70 Positive	Mossinger 1976 ⁹²	Cramps (legs)	13 Negative
Chevreil <i>et al</i> 1984 ⁸⁹	Ileus	58 Positive	Khan and Rawal 1976 ⁹³	Verruca plantaris	0 Positive
Valero 1981 ⁹⁰	Ileus	50 Positive			
Estrangin 1979 ⁹¹	Ileus	48 Negative			
Casteln 1979 ⁹²	Ileus	20 Positive			

*Comparison of homoeopathic treatments.

Although the number of trials is impressive, many questions remain. Virtually no evidence exists about the correct choice of the remedy and the potency to be used (different potencies or homeopathic substances should be compared in controlled trials). Hahnemann's principles have been brought into practice in innumerable ways, as is indicated by the differences among the trials presented here. The process of producing preparations (the percentage of alcohol in the solution; the number of times that the substance must be shaken during potentiation, etc) and their composition (especially when herbs are used) differ greatly among manufacturers. Also, there is no plausible explanation of the mechanisms through which homeopathy would act. Substances that contain only the solvent can have no pharmacological actions according to our present knowledge of physics and chemistry. If a homeopath is asked his or her opinion about these mechanisms, the most likely reply is "I do not know." In practice, if a treatment works knowledge of the mechanisms of action is not necessary, and numerous examples from regular medicine can be cited in which the mechanisms are hardly understood or not at all. However, to assume that an infinitesimally diluted substance in an alcoholic solution has pharmacological effects would mean that essential concepts of modern physics would have to be dismissed.

An important problem in reviewing the literature is publication bias. Especially with a controversial subject such as homeopathy, several problems may exist. More trials with positive results might have been submitted and accepted by "alternative" journals, whereas small trials with negative results might not have been submitted or might have been rejected. On the other hand trials with positive results might have been rejected and negative trials more readily accepted by "regular" journals. About one third of the trials were published in each of regular journals, alternative journals, and by other means of communication (proceedings, reports, dissertations, books). No relation between the result and the place of publication was seen. Negative results were reported in alternative journals 12 times, in regular journals seven times, and in other publications five times. When talking to authors of trials we identified at least six trials for which no manuscript had been submitted for publication. It is difficult to discover the true reasons for failure to submit an article for publication, but we think that the (possibly negative) results may have been an important factor in these cases.

Nevertheless, much evidence is available. We tried to decrease the effects of publication bias by extensively checking every possible source for publications or reports of trials. We wrote to many researchers and also visited several of them to learn whether there were any unpublished trials and to get further details of the published ones. We used strict criteria to select the best trials and based our main conclusions on the results of these. The amount of positive evidence even among the best studies came as a surprise to us. Based on this evidence we would be ready to accept that homeopathy can be efficacious, if only the mechanism of action were more plausible. The way in which the belief of people changes after the presentation of empirical evidence depends on their prior beliefs and on the quality of the evidence.^{10a,10b} Critical people who did not believe in the efficacy of homeopathy before reading the evidence presented here probably will still not be convinced; people who were more ambivalent in advance will perhaps have a more optimistic view now, whereas people who already believed in the efficacy of homeopathy might at this moment be almost certain that homeopathy works.

A trial of very high quality was that of the Groupe de Recherches et d'Essais Cliniques en Homéopathie,

initiated by the French Ministry for Social Affairs and performed by a group consisting of regular and homeopathic researchers.^{11,12} After the earlier publication of several trials in which homeopathy was shown to decrease the time to recovery of bowel movements after abdominal surgery, this hypothesis was retested in a rigorous trial comparing four groups of 150 patients (two groups were treated with opium C15 and raphanus C5, one group with indistinguishable placebo, and one group was not treated). No differences at all were found. Will more of such trials for other indications show the same results and refute the existing evidence?

The weight of the presented evidence will probably not be sufficient for most people to decide definitely one way or the other. The question arises, What further evidence would be needed? Investigations in animal or plant models may increase the belief of sceptical people before they have read the evidence from clinical trials, but if no positive results are found homeopaths may claim that homeopathy only works in humans. We did not assess the evidence from such investigations; Scofield concluded in 1984 in a comprehensive review article that "despite the great deal of experimental and clinical work there is only little evidence to suggest that homeopathy is effective. This is because of bad design, execution, reporting or failure to repeat experimental work."^{10c} If more (well performed) controlled trials in humans are demanded, cooperation between sceptical investigators and homeopaths is likely to make the trial results more convincing for many readers. The question is how many of such trials would be needed to draw definitive conclusions? The evidence presented in this review would probably be sufficient for establishing homeopathy as a regular treatment for certain indications. There is no reason to believe that the influence of publication bias, data massage, bad methodology, and so on is much less in conventional medicine, and the financial interests for regular pharmaceutical companies are many times greater. Are the results of randomised double blind trials convincing only if there is a plausible mechanism of action? Are review articles of the clinical evidence only convincing if there is a plausible mechanism of action? Or is this a special case because the mechanisms are unknown or implausible?

In our opinion, additional evidence must consist of a few well performed controlled trials in humans with large numbers of participants under rigorous double blind conditions. The results of the trials published so far, and the large scale on which homeopathy is brought into practice, makes such efforts legitimate.

This work was financed by a grant of the Dutch Ministry of Welfare, Public Health, and Cultural Affairs (Project No 87-35). We are grateful to Catherine Hill and Françoise Doyon of Institut Gustave Roussy, Villejuif, France; to David Taylor Reilly and Morag Taylor of the Glasgow Royal Infirmary for helpful discussions; and to many other researchers for discussions and help in obtaining published trials.

- Knipschild P, Kleijnen J, Riet ter G. Belief in the efficacy of alternative medicine among general practitioners in the Netherlands. *Soc Sci Med* 1990;31:625-6.
- Reilly DT, Taylor MA, McSharry C, Atkinson T. Is homeopathy a placebo response? Controlled trial of homeopathic potency, with pollen in hayfever as model. *Lancet* 1986;ii:881-6.
- Aulas JJ. *L'homéopathie*. Paris: Editions médicales Roland Bettez, 1985.
- Poitevin B. *Le devenir de l'homéopathie*. Paris: Doan Editeurs, 1987.
- Righetti M. *Forschung in der Homöopathie*. Göttingen: Ulrich Burgdorf, 1988.
- De Lange-de Klerk ESM, Feenstra L, Bezemer PD. Effectiviteitsonderzoek van homeopathische therapie bij kinderen met recidiverende bovenste luchtweginfecties. *Similia Similibus Curentur* 1986;16:78-82.
- Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;335:149-53.
- Ferley JJ, Zentgraf D, d'Almeida D, Balducci F. A controlled evaluation of a homeopathic preparation in the treatment of influenza-like syndromes. *Br J Clin Pharmacol* 1989;27:329-35.
- Reilly DT, Taylor MA, Campbell J, et al. Is homeopathy a placebo

- response? A controlled trial of homeopathic immunotherapy (HIT) in asthmatic patients [Abstract]. *Proceedings of the 45th Congress of the Liga Medicorum Homeopathica Internationalis* 1990.
- 10 McInerney CL. *Clinical trials*. New York: Oxford University Press, 1986. (Monographs in epidemiology and biostatistics. Vol 8.)
 - 11 Mavauz MJ, Guibard-Moscatto ML, Schwartz D, et al. Controlled clinical trial of homeopathy in postoperative ileus. *Lancet* 1988;1:528-9.
 - 12 GRECHO (Groupe de Recherches et d'Essais Cliniques en Homeopathie). Evaluation de deux produits homéopathiques sur la reprise du transit après chirurgie digestive. Un essai contrôlé multicentrique. *Presse Med* 1989;18:59-62.
 - 13 Wiesnauer M, Gaus W. Double-blind trial comparing the effectiveness of the homeopathic preparation Galphimia potentissima D6, Galphimia dilution 10° and placebo on pollinosis. *Arzneimittelforschung* 1985;35:1745-7.
 - 14 Arnal-Lasserre MN. Préparation à l'accouchement par homéopathie: expérimentation en double insu versus placebo [Dissertation]. Paris: Académie de Paris, Université René Descartes, 1986.
 - 15 Wiesnauer M, Gaus W. Wirksamkeitsvergleich verschiedener Potenzierungen des homöopathischen Arzneimitteils Galphimia glauca beim Heuschnupfen-Syndrom. Eine multizentrische, kontrollierte, randomisierte Doppelblindstudie. *Deutsche Apotheker Zeitung* 1986;126:2179-85.
 - 16 Zell J, Conner WD, Mau J, Feuerstein G. Behandlung von akuten Sprunggelenksdistorsionen. Doppelblindstudie zum Wirksamkeitsnachweis eines homöopathischen Salbenpräparats. *Fortschr Med* 1988;106:96-100.
 - 17 Valero EM. Etude de l'action préventive de Raphanus sativus 7 CH sur le temps de reprise du transit intestinal post-opératoire (à propos de 80 cas) - Pyroxygène 7 CH sur les infections post-opératoires (à propos de 128 cas) [Dissertation]. Grenoble: Université Scientifique et Médicale, 1981.
 - 18 Auglainier G. Action d'un traitement homéopathique sur la reprise du transit post opératoire. *Homéopathie* 1985;No 6:42-5.
 - 19 Wiesnauer M, Häussler S, Gaus W. Pollinosis-Therapie mit Galphimia glauca. *Fortschr Med* 1983;101:811-4.
 - 20 Bordes LR, Dorfman P. Evaluation de l'activité antitussive du sirop Drosuex: étude en double aveugle versus placebo. *Cahiers d'Otorhinolaryngologie* 1986;21:731-4.
 - 21 Feltz JP, Poutignat N, Azopardy Y, Charrel M, Zmirou D. Evaluation en médecine ambulatoire de l'efficacité d'un complexe homéopathique dans la prévention de la grippe et des syndromes grippaux. *Immunologie Médicale* 1987;No 20:22-8.
 - 22 Brigo B. Le traitement homéopathique de la migraine: une étude de 60 cas, contrôlée en double aveugle. *Journal of Liga Medicorum Homeopathica Internationalis* 1987;18:25.
 - 23 Maivall L, Weinfurter T, Mau J, Conner WD. Therapie des grippalen Infekts mit einem homöopathischen Kombinationspräparat im Vergleich zu Acetylsalicylsäure. Kontrollierte, randomisierte Einblindstudie. *Arzneimittelforschung* 1988;38:578-82.
 - 24 Wiesnauer M, Gaus W, Bohnacker U, Häussler S. Wirksamkeitsprüfung von homöopathischen Kombinationspräparaten bei Sinusitis. Ergebnisse einer randomisierten Doppelblindstudie unter Praxisbedingungen. *Arzneimittelforschung* 1989;39:620-5.
 - 25 Bignamini M, Bertoli A, Consolandi AM, et al. Controlled double-blind trial with Baryta carbonica 15 CH versus placebo in a group of hypertensive subjects confined to bed in two old people's homes. *British Homeopathic Journal* 1987;76:114-9.
 - 26 Cherrel JP, Soglier J, Destable MD. Reprise du transit intestinal en chirurgie digestive. Action homéopathique de l'Opium. *Presse Med* 1984;13:833.
 - 27 Gassinger CA, Wünsel G, Netter P. Klinische Prüfung zum Nachweis der therapeutischen Wirksamkeit des homöopathischen Arzneimittels Eupatorium perfoliatum D 2 (Wasserhanf composite) bei der Diagnose "Grippler Infekt". *Arzneimittelforschung* 1981;31:332-6.
 - 28 Rüter H. Ein homöopathischer doppelter Blindversuch zur Spezifität. *Hypothesis* 1986;No 12:473-6.
 - 29 Wiesnauer M, Gaus W. Orthostatische Dysregulation. Kontrollierter Wirkungsvergleich zwischen Euflexin 5 mg und dem homöopathischen Arzneimittel Haploppappus D2. *Zeitschrift für Allgemeinmedizin* 1987;63:18-23.
 - 30 Lewth G, Brown PK, Tyrrell DAJ. Controlled study of the effects of a homeopathic dilution of influenza vaccine on antibody titres in man. *Complementary Medical Research* 1989;3:22-4.
 - 31 Savage RH. A double blind trial to assess the benefit of Arica montana in acute stroke illness. *British Homeopathic Journal* 1977;66:107-20.
 - 32 Audreud LEC, Ara E, da Silva MSM, Cassaro A. Randomised double blind trial with homeopathy and placebo on rheumatoid arthritis. Sao Paulo, Brazil: Escola Paulista de Medicina, 1988.
 - 33 Skalioudis S, Hatzikostas H, Lambropoulou N, Othman A, Diamantidis S. Comparative clinical study of homeopathic and allopathic treatment in diabetes mellitus type II. *Proceedings of the 43rd Congress of the Liga Medicorum Homeopathica Internationalis* (Athens) 1988;549-56.
 - 34 Hitzberger G, Koen A, Dorici M, Bauer P, Wohlzogen FX. Kontrollierte randomisierte doppelblinde Studie zum Vergleich einer Behandlung von Patienten mit essentieller Hypertonie mit homöopathischen und pharmakologisch wirksamen Medikamenten. *Wien Klin Wochenschr* 1982;94:565-70.
 - 35 Carlini EA, Braz S, Troncone LRP, et al. Efeito hipnótico de medicação homeopática e do placebo. Avaliação pela técnica de "duplo-cego" e "cruzamento". *Rev. Ar. Med. Brasil* 1987;33:83-8.
 - 36 Gibson RG, Gibson SLM, MacNeill AD, Watson Buchanan W. Homeopathic therapy in rheumatoid arthritis: evaluation by double-blind clinical therapeutic trial. *Br J Clin Pharmacol* 1980;9:453-9.
 - 37 Fisher P. An experimental double-blind clinical trial method in homeopathy. Use of a limited range of remedies to treat fibrositis. *British Homeopathic Journal* 1986;75:142-7.
 - 38 Tsakopoulos I, Lahropoulou N, Hadjicostas C, Skalioudis S, Diamantidis S. Comparative study of homeopathic and allopathic treatment of benign paroxysmal positional vertigo. *Proceedings of the 43rd Congress of the Liga Medicorum Homeopathica Internationalis* (Athens) 1988;94-7.
 - 39 Hadjicostas C, Papis A, Drenou P, Papakonstantinou G, Diamantidis S. Comparative clinical study of homeopathic and allopathic treatment of haemorrhage of the upper digestive tract. *Proceedings of the 43rd Congress of the Liga Medicorum Homeopathica Internationalis* (Athens) 1988;536-41.
 - 40 Owen D. An intervention into the homeopathic treatment of patients with irritable bowel syndrome. *Wunderkreis: Congress of the Faculty of Homeopathy*, 1990.
 - 41 Gibson RG, Gibson SLM, MacNeill AD, et al. Salicylates and homeopathy in rheumatoid arthritis: preliminary observations. *Br J Clin Pharmacol* 1978;6:391-5.
 - 42 Master FJ. A study of homeopathic drugs in essential hypertension. *British Homeopathic Journal* 1987;76:120-1.
 - 43 Master FJ. Scope of homeopathic drugs in treatment of Broca's aphasia in a double blind trial. *Proceedings of the 42nd Congress of the Liga Medicorum Homeopathica Internationalis* (Arborea) 1987;330-4.
 - 44 Chakravarty BN, Sen JP, Mitra SK, et al. The effect of homeopathic drugs in tonsillitis. *Proceedings of the 32nd Congress of the Liga Medicorum Homeopathica Internationalis* (New Delhi) 1977;suppl:41-8.
 - 45 Gauthier JE. Essai thérapeutique comparatif de l'action de la chonidine et du Lachesis mutus dans le traitement des bouffées de chaleur de la ménopause [Dissertation]. Bordeaux: Université de Bordeaux II, 1983.
 - 46 Savage RH, Roe PF. A further double blind trial to assess the benefit of Arica montana in acute stroke illness. *British Homeopathic Journal* 1978;67:210-22.
 - 47 Dorfman P, Amoldo C, Riccio F, Tetat M, Verous G. Evaluation de l'activité d'Arica 5 CH. *Cahiers de Biochimie* 1988;No 98:37-82.
 - 48 Lecoq P. L.52. Les voies thérapeutiques des syndromes grippaux. *Cahiers de Biochimie* 1985;No 87:65-73.
 - 49 Lewis D. Double blind controlled trial in the treatment of whooping cough using drosuex. *Midlands Homeopathic Research Group, Research Newsletter* 1984;No 11:49-58.
 - 50 Schmidt W. Zur Therapie der chronischen Bronchitis. *Therapeutische Umschau* 1987;37:2803-9.
 - 51 Mössinger P. Zur Behandlung der Otitis media mit Pulsatilla. *Der Kinderarzt* 1955;16:531-2.
 - 52 Darns AE. Clinical investigation into the action of potencies. *British Homeopathic Journal* 1971;60:36-41.
 - 53 Mössinger P. Die Behandlung der Pharyngitis mit Physalis. *Allgemeine Homöopathische Zeitung* 1973;218:111-21.
 - 54 Mössinger P. Untersuchung zur Behandlung des akuten Fließschnupfens mit Euphorbia D3. *Allgemeine Homöopathische Zeitung* 1982;227:89-95.
 - 55 Hours P. Tentative d'appréciation de l'efficacité de l'homéopathie [Dissertation]. Pitié-Salpêtrière: Université Pierre et Marie Curie, 1982.
 - 56 Mössinger P. Untersuchung über die Behandlung der akuten Pharyngitis mit Phytolacca D2. *Allgemeine Homöopathische Zeitung* 1976;221:177-83.
 - 57 Masciello E, Feleci E. Dilutions de matériel, a pourcentage élevé de ADN et ARN, dans la prévention des virus épidémiques. *Proceedings of the 40th Congress of the Liga Medicorum Homeopathica Internationalis* (Lyon) 1985;271-4.
 - 58 Bungeztzani G. The results obtained by the homeopathic dilution (15CH) of an anti-influenza (anti-flu) vaccine. *Proceedings of the 43rd Congress of the Liga Medicorum Homeopathica Internationalis* (Athens) 1988;143.
 - 59 Ustianowski PA. A clinical trial of Strychnos in postcoital cystitis. *British Homeopathic Journal* 1974;63:276-7.
 - 60 Mössinger P. Zur therapeutischen Wirksamkeit von Hepar sulfuris calcareum D 4 bei Pyodermiten und Furunkeln. *Allgemeine Homöopathische Zeitung* 1980;225:22-7.
 - 61 Subramanyam VR, Mishra N, Rai Y, Rakshit G, Pattnaik NM. Homeopathic treatment of filariasis. Experience in an Indian rural setting. *British Homeopathic Journal* 1990;79:157-60.
 - 62 Carey H. Double blind clinical trial of Borax and Candida in the treatment of vaginal discharge. *Communications of the British Homeopathic Research Group* 1986;15:12-4.
 - 63 Castro D, Nogueira GG. Use of the nosode meningococcum as a preventive against meningitis. *Journal of the American Institute of Homeopathy* 1975;211-9.
 - 64 Reilly DT, Taylor MA. Potent placebo or potencies? A proposed study model with initial findings using homeopathically prepared pollens in hayfever. *British Homeopathic Journal* 1985;74:65-75.
 - 65 Rahlf VW, Mössinger P. Ass foetida bei Colon irritabile. Doppelblindversuch. *Dtsch Med Wochenschr* 1979;104:140.
 - 66 Rahlf VW, Mössinger P. Zur Behandlung des Colon irritabile. *Arzneimittelforschung* 1976;26:2230-4.
 - 67 Mössinger P. Misslungene Wirksamkeitsnachweise. *Allgemeine Homöopathische Zeitung* 1976;221:26-31.
 - 68 Mössinger P. Zur therapeutischen Wirksamkeit von Absinthum bei der Cholezystopathie. In: Mössinger P, ed. *Der praktische Arzt als Fachmann für Erfahrung und Beobachtung*. Heidelberg: Karl F Haug, 1974:99-101.
 - 69 Estrangin M. Essai d'approche expérimentale de la thérapeutique homéopathique [Dissertation]. Grenoble: Université Scientifique et Médicale de Grenoble, 1979.
 - 70 Castelin T. Etude de l'action homéopathique de Raphanus sativus niger 5 CH et d'Opium 15 CH sur la reprise du transit en chirurgie digestive post-opératoire [Dissertation]. Bordeaux: Université de Bordeaux II, 1979.
 - 71 Shapley AB, Berry H, Brooker G, et al. Controlled trial of homeopathic treatment of osteoarthritis. *Lancet* 1983;1:97-8.
 - 72 Finber P, Greenwood A, Huxkinson EC, Turner P, Belon P. Effect of homeopathic treatment on fibrositis (primary fibromyalgia). *BMJ* 1989;299:365-6.
 - 73 Bourgeois JC. Protection du capital veineux chez les perfusés au long cours dans le cancer du sein. Essai clinique en double aveugle: Arica contre placebo [Dissertation]. Bobigny: Université Paris Nord, 1984.
 - 74 Casanova PA. Essai clinique d'un produit appelé "Urturbine". Metz: Laboratoires Lehnring, 1981.
 - 75 Pinsten RJFH, Baker GPI, Joes G, Davey RW, Jonsk S. Does Arica reduce pain and bleeding after dental extraction? A placebo controlled pilot study conducted by the Midland Homeopathy Research Group (MHRG) in 1980/81. *Communications of the British Homeopathic Research Group* 1986;No 15:3-11.
 - 76 Bertier P. Etude sur 80 cas en patiente privée d'une prémédication homéopathique pour les extractions et la chirurgie buccale. *Proceedings of the 40th Congress of the Liga Medicorum Homeopathica Internationalis* (Lyon) 1985;79-82.
 - 77 Albertini H, Goldberg W, Sangov, Tourla. Bilan de 60 observations dentaires. Hypericum-Arrica contre placebo dans les névralgies dentaires. *Homéopathie Française* 1984;71:47-9.
 - 78 Campbell A. Two pilot controlled trials of Arica montana. *British Homeopathic Journal* 1976;65:154-8.
 - 79 Hildebrandt G, Ekze C. Über die Wirksamkeit einer B-handlung des Muskelkaters mit Rhus toxicodendron D4. In: *Hufelandgesellschaft für Gesamtmedizin. Wissenschaftliches Archiv der Hufelandgesellschaft für Gesamtmedizin*. Vol 1. Heidelberg: Karl F Haug, 1983.
 - 80 Leaman AM, Gorman D. Cantkars in the early treatment of minor burns. *Arch Emerg Med* 1989;6:259-61.
 - 81 Geiger G. Klinische Erfahrungen mit Traumaed bei Weichteiltraumaten und Frakturen und mit Vertugol bei der Commotio cerebri acuta. *Medizinische Welt* 1968;No 18:1203-4.

- 82 Kubista E, Möller G, Spöna J. Behandlung der Mastopathie mit cyclische Mastodynie: klinische Ergebnisse und Hormonprofile. *Gynäkologische Rundschau* 1986;26:65-79.
- 83 Michaud J. Action d'Apis mellifica et d'Arnica montana dans la prévention des oedèmes post-opératoires en chirurgie maxillo-faciale à propos d'une expérimentation clinique sur 60 observations [Dissertation]. Nantes: Université de Nantes, 1981.
- 84 Alergen H. Therapie posttraumatischer Schwellungen mit Traumcel. Beitrag zur Relation "Dosis/Wirkung" eines Kombinationspräparates. *Münchener Medizinische Wochenschrift* 1969;111:298-300.
- 85 Caspar J, Foerstel G. Traumcel bei traumatischen Weichteilschwellungen. *Therapeutische Umschau* 1967;17:892-5.
- 86 Khan MT. Clinical trials for hallux valgus with the Marigold preparations. *Proceedings of the 40th Congress of the Liga Medicorum Homoeopathica Internationalis* (Lyon) 1985:233-5.
- 87 Anonymous. The Cantharis study. *Midlands Homoeopathic Research Group, Research Newsletter* 1980;No 3:19-20.
- 88 Delaunay M. Homéopathie, la maternelle. *Médecines Douces* 1985;Sep:35-7.
- 89 Heulluy B. Essai randomisé ouvert de L. 72 (spécialité homéopathique) contre diazepam 2 dans les états anxiodépresseurs. Metz: Laboratoires Lehning, 1985.
- 90 Ponti M. Evaluation d'un traitement homéopathique du mal des transports. Bilan de 93 observations. In: Boiron J, ed. *Recherches en homéopathie*. Sainte-Foy-les-Lyon: Fondation Française pour la Recherche en Homéopathie, 1986:71-4.
- 91 Vu Din Sao, Delaunay M. Médecine douce et sport dur: un mariage heureux. *Homéopathie Française* 1983;71:147-9.
- 92 Dexpert M. Prévention des naupathies par Cocculine. *Homéopathie Française* 1987;75:353-5.
- 93 Aibeau JP, Jobert J. Aconit en dilution homéopathique et agitation post-opératoire de l'enfant. *Pédiatrie* 1990;45:463-6.
- 94 Davies AE. A pilot study to measure aluminium levels in hair samples of patients with dementia and the influence of aluminium 30c compared with placebo. *Midlands Homoeopathic Research Group, Research Newsletter* 1988;No 18:42-6.
- 95 Coudert-Deguillem M. Etude expérimentale de l'action du Caulophyllum dans le faux travail et la dystocie de démarrage [Dissertation]. Limoges: Université de Limoges, 1981.
- 96 Kennedy CO. A controlled trial. *British Homoeopathic Journal* 1971;60:120-7.
- 97 Paterson J. Report on mustard gas experiments. *British Homoeopathic Journal* 1943;23:131-42.
- 98 Basu TK. Studies on the role of physostigma venenosum in the improvement of simple myopia. *Hahnemannian Gleamings* 1980;47:224-31.
- 99 Harivaux E. La recherche clinique à l'Institut Boiron. *Homéopathie* 1987;No 5:53-8.
- 100 Kirchhoff HW. Ein klinischer Beitrag zur Behandlung des Lymphödems. *Der Praktische Arzt* 1982;21:621-33.
- 101 Kienle G. Wirkung von Carbo Betulae D 6 bei respiratorischer Partialinsuffizienz. *Arzneimittelforschung* 1973;23:840-2.
- 102 Ventoskovskiy BM, Popov AV. Homoeopathy as a practical alternative to traditional obstetric methods. *British Homoeopathic Journal* 1990;79:201-5.
- 103 Schwab G. Lässt sich eine Wirkung homöopathischer Hochpotenzen nachweisen? Eine kontrollierte Cross-over Doppelblindstudie bei Hautkrankheiten. *Allgemeine Homöopathische Zeitung* 1990;235:135-9.
- 104 Khan MT, Rawal RS. Comparative treatments of verruca plantaris. *Proceedings of the 31st Congress of the Liga Medicorum Homoeopathica Internationalis* (Athens) 1976:265-9.
- 105 Knipschild P. Changing belief in iridology after an empirical study. *BMJ* 1989;299:491-2.
- 106 Knipschild P, Leffers P. De informatiewaarde van empirisch onderzoek. *Tijdschrift voor Sociale Gezondheidszorg* 1990;68(suppl):31-2.
- 107 Scofield AM. Homoeopathy and its potential role in agriculture. A critical review. *Biological Agriculture and Horticulture* 1984;2:1-50.

(Accepted 4 December 1990)

Trials of homoeopathy

SIR, — We would like to reply to several letters that have commented on our meta-analysis of clinical trials of homoeopathy.¹

Dr Peter Fisher and colleagues address the problem of evaluating research for which the report does not include full details of the characteristics of the patients or methodology, etc.² We believe that sufficient details must be present to enable the reader to judge the evidence. It would surely be very complicated if every reader had to go after the necessary information. They say that the journal did not allow them to put more information in their report of homoeopathic treatment. A subsequent letter,³ however, did not contain information that would increase our methodological score, nor did a duplicate publication of the same trial.⁴

Dr Fisher and colleagues also oppose our view on crossover trials. We have doubts about the prognostic comparability at the start of the second period of these trials unless the treatment does not work. This will especially be the case when subjective symptoms are most important. If the results of only the first period are valid they must be presented separately. One of the goals of homoeopathy is long lasting treatment effects, preferably after short treatment. Therefore we think that crossover designs (including studies of one patient) do not make much sense despite the fewer patients needed. In the extreme, the optimum number of patients in such trials is zero.

In checks on blinding (before treatment effects are expected) the subjects' answers must be compared between the treatment and control groups. Breaking of blinding is suspected if there are differences in the distribution of the answers between the groups. The influence of leading questions might be a problem only if almost all the patients answered "don't know." In practice many patients will guess that they are receiving placebo or active treatment.⁵

Dr Marc Girard rightly states that the evidence available is less impressive if each homoeopathic remedy is considered separately for each new medication.⁶ Before assessing the effects of all

these therapies and variations of homoeopathy it should be established whether homoeopathy has a specific effect in at least some cases. It is not surprising that opinions differ over whether the existing evidence is sufficient for such efforts, which is illustrated by the remarks of Dr Michael Baum⁷ and Sir Michael Drury and colleagues.⁸

We had extensive discussions about how further evidence should be obtained. The results of our review would probably be interpreted differently if laboratory studies showed convincing evidence that there is some action of high potencies. Two of us (PK, GtR) thought that this kind of evidence has priority. On the other hand, if laboratory studies show no effects homoeopaths will say that homoeopathy works only in humans. In that case only rigorously performed randomised trials will confirm or refute the existing state of affairs (JK). Meanwhile, reports of new controlled trials are being published.

JOS KLEIJNEN
PAUL KNIPSCHILD
GERBEN TER RIET

Department of Epidemiology and
Health Care Research,
University of Limburg,
PO Box 616, 6200 MD Maastricht,
The Netherlands

- 1 Kleijnen J, Knipschild P, ter Riet G. Clinical trials of homoeopathy. *BMJ* 1991;302:316-23. (9 February.)
- 2 Fisher P, Huskisson EC, Scott DL, Turner P, Whitmarsh TE. Trials of homoeopathy. *BMJ* 1991;302:727. (23 March.)
- 3 Fisher P, Huskisson EC, Turner P, Belon P. Complementary medicine. *BMJ* 1989;299:1401-2.
- 4 Fisher P. Rus toxicodendron in the treatment of fibromyalgia: a double-blind, placebo-controlled trial, with cross-over. *Journal de la Organisation Médicale Homéopatique Internationale* 1988;1(3):26-8.
- 5 Kessels AGH, Cardynals RLLM, Bonger RLL, et al. The effectiveness of the hair-restorer "Dabao" in males with alopecia androgenetica. A clinical experiment. *J Clin Epidemiol* 1991;44:439-47.
- 6 Girard M. Trials of homoeopathy. *BMJ* 1991;302:727. (23 March.)
- 7 Baum M. Trials of homoeopathy. *BMJ* 1991;302:529. (2 March.)
- 8 Drury M, Greenfield SM, Hobbs FDR, Kenkre J, Salter R. Trials of homoeopathy. *BMJ* 1991;302:529. (2 March.)
- 9 Wiesener M, Gaus W. Wirksamkeitsnachweis eines Homöopathikums bei chronischer Polyarthrit. Eine randomisierte Doppelblindstudie bei niedergelassenen Ärzten. *Aktuelle Rheumatologie* 1991;16:1-9.

CHAPTER 13

REVIEW ARTICLES AND PUBLICATION BIAS*

Jos Kleijnen, Paul Knipschild

* Submitted for publication

Summary

Publication bias occurs if the results from studies which have not been published, are different from the published ones. From a Bayesian viewpoint, it also concerns non-publication of studies with similar results as the published ones, because the strength of the evidence will be influenced. Publication bias complicates the interpretation of reviews and meta-analyses. If favourable results are published more often, there will be an over-estimation of the effects of a treatment.

There have been several attempts to assess the magnitude of publication bias. Unpublished trials could be identified by means of a survey among researchers and the results could subsequently be compared with the outcomes of published trials. Also, the results from published trials could be compared with trials from a registry. Furthermore, the results from registered but unpublished trials could be compared with those of registered and subsequently published trials. Studies addressing publication bias have shown that it is a serious problem, which complicates the interpretation of reviews.

In assessments of publication bias other factors must be taken into account. These include the mode of publication: refereed journals, other journals, books etcetera. Differences could also be related to the quality of trials. Finally, the source of funding may influence both the results and subsequent publication.

Publication bias can only be avoided by registration of all trials before data collection is started; several of such registries have already been installed. Perhaps, if more of such registries exist, reviewers could only use registered trials for their main conclusions. All other information could then be considered sensitive to publication bias.

Introduction

The number of articles reporting research findings in the medical literature is overwhelming. Because it is hardly possible to keep track of all the new information, there is an increasing demand for overview articles summarizing the state of affairs on specific subjects. The value of these reviews depends, among other factors, on how completely and how validly the included publications represent the total available evidence. A major problem in summarizing the existing literature is finding all publications addressing the subject of interest. One might restrict a review to full articles published in refereed journals indexed in computerized systems for the retrieval of literature. However, valuable, high quality information might also be found in other journals, books, dissertations, abstracts and unpublished reports. It is very difficult and time consuming to search for information from these sources.

If all evidence can be included, there is the problem that the quality and reliability of the data is not the same in all publications. Even within refereed journals, the quality of published data varies considerably. There may also be differences between the average quality of publications in refereed journals compared with other journals, books, dissertations etcetera. The reported outcome may be associated with the quality, and then a review may be seriously biased if all these data are included and bear the same weight on the assessment of the evidence. Studies addressing this problem will be presented in this article. It seems logical that conclusions must be drawn using only evidence of high quality. Consequently, to separate the chaff from the wheat, the methodological quality of included studies must be assessed. Reviews in which such a methodological assessment of the quality of trials is emphasized, are called criteria-based meta-analyses. However, even if all of the published literature is identified and included in a review and if the methodological quality is taken into account, the problem remains that there might be relevant research findings that have not been published. This has been illustrated in a publication of Hemmeniki (1980), who studied reports of clinical trials included in applications by drug companies to licensing authorities in Finland and Sweden. Of these trials, about 50% had not been published or had only a summary available.

Publication bias

Publication bias occurs if the results from studies which have not been published, are different from the published ones (Begg & Berlin 1989 and Dickersin 1990). It also occurs that only part of the findings is published (for example after multiple testing) or, more subtly, that negative results of a trial are not separately published, but only appear in a meta-analysis from the same author or company, between many other positive findings. From a Bayesian outlook on the evidence from clinical trials, the appreciation of the evidence is influenced regardless of the outcome, at least if the total number of published trials is not very large. Bayesian inference works with subjectively anticipated probabilities, which are compared with empirical evidence, and then adjusted to posterior probabilities. Thus, more evidence, even if there is no difference in outcome, may lead to a greater change in belief in the hypothesis. From this viewpoint, unpublished trials may bias the assessment even if the outcome does not differ from published trials. Publication bias complicates the interpretation of reviews and meta-analyses. If favourable results are published more often, there will be an overestimation of the effects of a treatment. In a review of homoeopathy 77% of 105 controlled trials showed positive effects (Kleijnen et al 1991). This indicates a very large impact of bias, including publication bias, if one assumes it impossible that extreme dilutions have any pharmacological effect.

It is well known that articles reporting positive findings have a greater chance of being accepted for publication in medical journals compared with those reporting negative results (Sterling 1959, Greenwald 1975, Mahoney 1977, Atkinson 1982, Coursol & Wagner 1986, Koren 1989). But failing to write a report by the researcher is probably an even greater cause of bias (Dickersin 1987, Easterbrook et al 1991). There may be several causes for the latter. Something might have gone wrong during the investigation or the methodology used was considered not satisfactory in hindsight. Also, the results could have been disappointing and interest in writing them down disappeared. A combination of various reasons will also occur.

The goal of this review article is to stress the importance of publication bias. Attention will also be paid to factors influencing the likelihood of publication, and to methodological problems when measuring its impact.

Assessments of publication bias

The extent to which publication bias occurs is difficult to assess and will probably depend on the subject at hand. There have been several attempts to assess the magnitude of publication bias. These assessments must contain a comparison of the results of published and unpublished trials, in order to establish whether there is any bias. From a methodological viewpoint two approaches are commonly used. Firstly, unpublished trials are identified by means of a survey among researchers and subsequently the results are contrasted to the results of published trials. Secondly, a comparison could be made of the results from published trials versus registered trials. Examples of registries are projects approved by ethical committees, or research data banks e.g. for cancer trials, in which trials are registered in an early stage, before data are collected. In this second approach, also trials within registries (published versus unpublished) could be compared. Articles on publication bias were found by a MEDLINE search with the keyword *Publication bias* which is in use since 1983. The first hit however was an article published in 1985 (Begg 1985). Further articles were found by extensively checking references from articles we already possessed, and by personal communication with researchers in this field. We can only hope that there is no publication bias among the articles dealing with this very problem.

Surveys for unpublished trials

One possible approach to identify unpublished studies is a survey among researchers. Apparent limitations are low response rates and, very likely, unrepresentative samples of all unpublished trials will be obtained. Inherently, only limited conclusions can be drawn from such surveys.

Sommer (1987) identified studies addressing several topics in menstrual cycle research after sending a questionnaire to members of the Society for Menstrual Cycle Research residing in the United States and Canada. Twenty-two out of 30 studies (73%) reporting positive results had been published, compared with 14 out of 26 studies (54%) without positive findings.

A survey among 318 authors of published trials (on different subjects, identified through the MEDLINE database and Current Contents) yielded 178 completed unpublished trials with a trend specified (Dickersin et al 1987). Of these 26 (14%) favoured the new therapy compared with 423 of 767 (55%) published reports. Major reasons for nonpublication were negative results and lack of interest. It also appeared that nonpublication was mainly a result of failure to write a manuscript and submit the trial rather than rejection of submitted papers.

Shadish et al (1989) performed a meta-analysis of 165 randomized studies on family/marital psychotherapy. In an attempt to identify additional unpublished studies they drew a random sample of 519 out of 14,000 members of associations for marital and family therapy, behavior therapy and psychology. They were asked if they had any file drawer studies. A total of 375 respondents yielded 3 such unpublished studies. It was estimated that in the whole population 112 (95% Confidence interval [CI]: 23-325) file drawer studies existed that would have qualified for inclusion in the meta-analysis.

Hetherington et al (1989) sent letters to 42,000 obstetricians and pediatricians in 18 countries in an attempt to obtain information about unpublished trials. They were notified of 395 unpublished randomized trials, of which only 18 had been completed more than 2 years before the survey, and 125 had ceased recruitment within the two years prior to the survey. The other trials were ongoing. They concluded that publication bias cannot not be addressed successfully by retrospective attempts to obtain information on unpublished trials.

Published versus registered trials

This model is based on the selection of trials registered in advance in a registry, and thus free from publication bias. A review of published clinical trials (identified by a literature search) is then compared with a review of all registered trials.

Simes (1986) compared the results from trials assessing the effects on survival for chemotherapies in advanced ovarian cancer and also in multiple myeloma. From trials registered in the International Cancer Research Data Bank a median survival ratio of the intervention and the control treatments was calculated. There was a trend for registered trials to show lower survival ratios. In advanced ovarian cancer the ratio was 1.16 for published trials and 1.06 for registered trials with data. In myeloma trials the ratios were 1.26 and 1.14.

Comparisons of published and unpublished trials within a registry

Also trials within a registry, which have been published could be compared with trials with data available (from the same registry) which have not been published. This is the most valid approach to assess publication bias, because there is no selection bias if all trials can be followed up.

487 Research projects approved by the Oxford Research Ethics Committee were studied for evidence of publication bias (Easterbrook et al 1991). Studies with statistically significant results were more likely to be published than trials showing no difference (odds ratio 2.23). An increased likelihood of publication was found to be associated with increasing sample size; this was also the case for studies with a high rating by the investigator of the importance of the study results. The tendency towards publication bias was greatest for observational and laboratory-based experimental studies (odds ratio 3.79) and was not clearly found for randomized clinical trials (odds ratio 0.84).

Factors influencing the interpretation of assessments of publication bias.

In these assessments other factors must be taken into account. We have already mentioned that there might be differences in outcomes between publications in refereed journals, other journals, books etcetera. Part of these differences might be related to the quality of trials; the quality may very well influence the results. Finally, the source of funding may influence both the results and the likelihood of subsequent publication.

Influence of the source of publication

Glass et al (1981) compared the results from meta-analyses on different subjects representing both published and unpublished literature. In all instances in which the comparison could be made, the average experimental effects from studies published in journals was larger than the corresponding effect estimated from theses and dissertations.

White (1982) found some 100 studies in which the relation between socioeconomic status and academic achievement was examined. He found different correlation coefficients when considering the type of publication: for books 0.51; journals 0.34; and unpublished 0.29. Also the quality of the evidence influenced the outcome: valid 0.30; fairly valid 0.36; invalid 0.49.

Goldman & Loscalzo (1980) found that from 276 identified cardiology abstracts only 50% led to full-length articles in peer reviewed journals. Meranze et al (1982) found that 30-44% of abstracts presented at several anaesthesia meetings were subsequently published. According to McCormick & Holmes (1985) about half of the abstracts presented at pediatric meetings were eventually published. Unfortunately, the outcomes of the abstracts and the articles were not compared in the above investigations.

Chalmers et al (1990) identified 176 summary reports in the Oxford Database of Perinatal Trials. More than two thirds of these reports were abstracts, the remainder letters to the editor and short reports. Full reports (of these summaries, which were subsequently published) were identified for 36% of the 176 summary reports. Trials judged to be of good quality and those containing positive interpretation of the results, were no more likely followed by full reports than others. Of trials reporting positive results 33% was subsequently published and, surprisingly, for trials reporting neutral / negative results this was 41%. The rate of subsequent publication was higher among trials with a sample size above the median (47 participants).

Influence of the quality of trials

Sacks et al (1983) reported that for six therapies 79% of trials using historical controls found a new therapy effective, while only 20% of randomised clinical trials of the same treatment confirmed these outcomes. It concerned studies into coronary artery surgery, anticoagulants for acute myocardial infarction, surgery for esophageal varices, fluorouracil for colon cancer, BCG for melanoma and diethylstilbestrol for habitual abortion.

Berlin et al (1989) examined a sample of published cancer clinical trials. Small studies showed large treatment effects and large studies relatively small effects. The differences observed were 41% for overall survival, 79% for disease-free survival and 17% for response rates.

Vandenbroucke (1988) reviewed studies of the effect of passive smoking on lung cancer. The relative risks were compared for studies with large and small sample size. These risks should scatter around the underlying true risk, with greater scatter among the smaller studies. It may be an indication of publication bias if small studies with negative results are lacking. This turned out to be the case for studies in men only.

Ter Riet et al (1989) analysed 91 controlled trials of acupuncture. The methodological quality of trials was assessed using criteria such as the number of patients included, randomization, blinding of the effect measurement, sensible and effect measurements and adequate data presentation. It turned out that with an increasing methodological quality more negative results were reported. The best trials all reported no difference between the index and control groups.

Kleijnen et al (1991) analysed 105 controlled trials of homoeopathy using similar criteria of quality. Most trials showed positive effects: 68% of the better trials and 80% of the lower quality trials.

In an effort to assess variation in study quality Begg et al (1987) used randomization and sample size to compare 218 trials in clinical cancer research identified using a MEDLINE search of July to December 1985. Randomized studies turned out to produce half as likely a positive conclusion.

In the study of Easterbrook et al (1991) the hypothesis that publication bias might be explained by confounding factors, especially the quality of study design, was also tested. They suggest that bias against the publication of null results is a powerful force in the publication process, regardless of quality.

Influence of the source of funding

In research sponsored by pharmaceutical companies or other commercially oriented sources, commercial reasons might influence the researcher. Davidson (1986) reported that the source of funding influenced the outcome of published clinical trials. From a total of 107 identified trials in 5 well known medical journals in 1984, 33 of 37 pharmaceutically supported trials (89%) and 43 of 70 generally supported trials (61%) showed positive findings. However, Easterbrook et al (1991) found that studies sponsored by a pharmaceutical company were less likely to be published, whatever the results.

Discussion and conclusions

In surveys for unpublished trials, it seems likely that if the outcome plays a role in the decision to write a manuscript, it will also influence the response of investigators to the survey. Any sample of new unpublished trials will again be biased in favour of positive results. There is also the problem of whom should be addressed? Evident candidates are authors of already published trials. However, people who did one study on a certain subject, but never published it, are not reached because nobody knows them. Furthermore, the response rate may be very low, and people might not be inclined to admit that they performed research that has not been published. Finally, if the number of identified unpublished trials is very low, this can lead to very crude estimations of the total amount of unpublished studies (see for instance Shadish 1989). The above points will probably influence the result in the same direction: towards an underestimation of publication bias. Consequently, the results from surveys may only be interesting if a substantial number of unpublished trials is identified.

Identification of published trials can be fairly complete if elaborate strategies are used. With the exception of comparisons within a group of registered trials, there is always only a selection of unpublished trials present. If also a selection of published trials is reviewed, matters become very complicated.

Thus, the most valid approach is comparing the results of a known cohort of trials, such as those approved by an ethical committee. If this cohort is followed in time, some trials are not performed or stopped, but others are completed and a manuscript is written. Selection bias will be limited if a cohort of registered trials is followed up because all trials can be identified. If sufficient relevant information can be obtained (for example through the principal investigators), a comparison of the likelihood of publication for trials reporting positive results and trials showing no positive results may give an impression of the amount of publication bias. The only drawback in these comparisons may be the trials not performed or stopped (comparable with drop-outs in a cohort study). If (negative) results, e.g. after interim analyses, were the cause of stopping the trial, publication bias is introduced.

Another limitation of this approach is that these registered trials are often limited to a certain specialism (e.g. pediatrics, oncology) or some other specific category, thus making extrapolation to other areas difficult. Abstracts from specified meetings could be con-

sidered as a cohort, and similar comparisons can be made as for trials from registries. However, abstracts must also be submitted and accepted for publication and a first selection has already taken place.

Whether the results of trials are published in a refereed journal and not somewhere else will very likely be influenced by a judgement of the quality by the author as well as the outcome of the trial. However, also other factors may influence the mode of publication. For instance, trials of alternative medicine showing positive evidence may very well be rejected by refereed regular journals, and the only possible way to publish the (sometimes high quality) evidence may be some other form of communication. There are even examples of Nobel prize winners of who the prime articles were initially rejected by scientific journals. In any case, the presented examples show that publication bias may go in any direction depending on the subject at hand when considering the mode of publication.

In comparisons of the outcomes of high quality versus low quality trials, one never knows whether differences are caused by publication bias or other biases which are more likely to occur when the quality of the trials decreases. The general picture seems to be that high quality trials tend to show more confirmations of the null hypothesis than trials or other evidence of lower quality. In review articles, the quality of the evidence should always be taken into account.

Finally, it might also be worthwhile to consider the source of funding. Part of the trials funded by the pharmaceutical industry remains unpublished, and the outcome is probably an important factor in the decision whether to publish or to file a report of a trial.

What remains is the question of how to make sure that publication bias does not occur. It is not difficult to find the answer, and it has been proposed many times before: registration of all trials before they are started. Several registries of controlled trials have already been installed, examples have been presented. Perhaps, if more of such registries exist, reviewers could restrict themselves to only registered trials for their main conclusions. Other information could then be viewed as most of nowadays reviews: prone to publication bias.

References

1. Atkinson DR, Furlong MJ, Wampold BE. Statistical significance, reviewer evaluations, and the scientific process: Is there a (statistical) significant relationship? *Journal of Counseling Psychology* 1982;29(2):189-94.
2. Begg C. A measure to aid in the interpretation of published clinical trials. *Stat Med* 1985;4:1-9.
3. Begg CB, Pocock SJ, Freedman L, Zelen M. State of the art in comparative cancer clinical trials. *Cancer* 1987;60:2811-5.
4. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst* 1989;81:107-15.
5. Berlin JA, Begg CB, Louis TA. An assessment of publication bias using a sample of published clinical trials. *J Am Stat Assoc* 1989;84:381-92.
6. Chalmers I, Adams M, Dickersin K, Hetherington J, Tarnow-Mordi W, Meinert C, Tonascia S, Chalmers TC. A cohort study of summary reports of controlled trials. *JAMA* 1990;263:1401-5.
7. Coursol A, Wagner EE. Effect of positive findings on submission and acceptance rates: a note on meta-analysis bias. *Professional Psychology: Research and Practice* 1986;17(2):136-7.
8. Davidson RA. Source of funding and outcome of clinical trials. *J Gen Intern Med* 1986;1:155-8.
9. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H. Publication bias and clinical trials. *Controlled Clinical Trials* 1987;8:343-53.
10. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;263:1385-9.
11. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;337:867-72.

12. Glass GV, McGaw B, Smith ML. Meta-analysis in social research. Sage Publications, Beverly Hills London 1981:64-7.
13. Goldman L, Loscalzo A. Fate of cardiology research originally published in abstract form. *N Engl J Med* 1980;303:255-9.
14. Greenwald AG. Consequences of prejudice against the null hypothesis. *Psychol Bull* 1975;82:1-20.
15. Hemminki E. Study of information submitted by drug companies to licensing authorities. *Br Med J* 1980;280:833-6.
16. Hetherington J, Dickersin K, Chalmers I, Meinert CL. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989;84:374-80.
17. Kleijnen J, Knipschild P, Riet ter G. Clinical trials of homoeopathy. *Br Med J* 1991;302:316-23.
18. Koren G, Graham K, Shear H, Einarson T. Bias against the null hypothesis: the reproductive hazards of cocaine. *Lancet* 1989;ii:1440-2.
19. Mahoney MJ. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research* 1977;1(2):161-75.
20. McCormick MC, Holmes JH. Publication of research presented at the pediatric meetings. Change in selection. *Am J Dis Child* 1985;139:122-6.
21. Meranze J, Ellison N, Greenhow E. Publications resulting from anesthesia meeting abstracts. *Anesth Analg* 1982;61:445-8.
22. Riet ter G, Kleijnen J, Knipschild P. Nawoord en aanbevelingen [De effectiviteit van acupunctuur]. *Huisarts en Wetenschap* 1989;32:308-12.
23. Sacks HS, Chalmers TC, Smith Jr H. Sensitivity and specificity of clinical trials. Randomized v historical controls. *Arch Intern Med* 1983;143:753-5.
24. Shadish WR, Doherty M, Montgomery LM. How many studies are in the file drawer? An estimate from the family/marital psychotherapy literature. *Clin Psychol Rev* 1989;9:589-603.
25. Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986;4:1529-41.
26. Sommer B. The file drawer effect and publication rates in menstrual cycle research. *Psychology of Women Quarterly* 1987;11:233-42.
27. Sterling T. Publication decisions and their possible effects on inferences drawn from tests of significance - or vice versa. *J Am Stat Assoc* 1959;54:30-4.
28. Vandenbroucke JP. Passive smoking and lung cancer: a publication bias? *Br Med J* 1988;296:391-2.
29. White KR. The relation between socioeconomic status and academic achievement. *Psychological Bull* 1982;91:461-81.

CHAPTER 14

EPILOGUE

This chapter describes some general problems and ideas which appeared while reviewing the literature and after the publication of the findings. Specific problems have already been dealt with in the discussion sections of the presented articles.

Problems and drawbacks in criteria-based meta-analysis

There are some obvious difficulties in evaluating research of which the publication does not include full details of patient characteristics, methodology etc. Sufficient details must be present to enable the reader to adequately judge the evidence. It would surely be very complicated if every single reader must go after the necessary information. We chose to evaluate all trials from a methodological viewpoint, using a criteria-based scoring system. Each criterion could only be scored if the information was actually mentioned in the publication. For instance, in a truly double-blind trial, randomization has probably been performed. However, the criterion randomization was only scored if it was specifically mentioned. This approach might lead to a lower quality assessment. On the other hand, one is never completely certain whether a double-blind trial is really double-blind, if blinding is not checked. Such a check was virtually never done in the trials presented. This, and similar discrepancies for other items between the information in the publication and what actually happened, may of course lead to a quality assessment which is not valid.

All trials have been scored at least twice. Differences were mainly caused by reading errors, and could easily be resolved by discussions. Large differences did not occur. Especially the scoring criteria of patient characteristics and the effect measurement must be carefully defined, because here appropriateness has to be judged from clinical experience. Also, the criteria and the weight of separate items may be different depending on the subject matter and on the endpoints used. For instance, our weight on large numbers of patients may be less if all trials included in a meta-analysis have been registered, and therefore the likelihood of publication bias is lower. Another reason for our emphasis on large numbers of participants is our worry about prognostic comparability at baseline. If small numbers of participants are included, randomization alone is not sufficient to ensure prognostic comparability at baseline. Prestratification (matching in the extreme) would help the luck of the draw. If most trials would start with a qualification period, this could also decrease the weight of the criterion of patient numbers. In such a qualification period, previous treatments can be washed out, accurate measurements of inclusion criteria and baseline characteristics obtained, compliance checked and bad compliers excluded, placebo responders identified and excluded etc. The efficiency of a trial could thus be enhanced, and fewer participants would be needed.¹ Also, the weight of the items might be different when hard endpoints (death) or soft endpoints (pain assessments) are used. Blinding might be less important in the former and extremely important in the latter case.

The criteria we have used are arbitrary, but well accepted. Moreover, the reader can check the evaluation, and apply different weights and/or different criteria to see whether substantial changes would occur. These criteria may give the reader an impression of why we think that a certain trial is well performed or not. The total score is only an indication of the quality of a trial. Differences of a few points are not important. In some of our reviews, we used a score of 60-65 points as a cut-off to call a trial reasonably well performed. A trial scoring 80 points or more might be considered a good trial, and the

best trials scored 90 points. A complicating factor is the so-called fatal mistake: one serious error which makes the whole trial useless. The likelihood of such a fatal mistake increases, of course, with an increasing number of methodological shortcomings.

For the assessment of the treatment effects, we have chosen to use the conclusions of the authors. Of course, it would be preferable if the reviewer could check the conclusions. Unfortunately this was frequently impossible, because the necessary data were lacking. A consequence of this approach is that in some cases (approx. 5%) the conclusions seemed to be too optimistic, i.e. the author mentioned positive results while we thought that the presented data did not support that conclusion. The opposite occurred seldomly.

Other authors have also used criteria for the assessment of methodological quality. Chalmers et al. proposed an index of quality of randomized clinical trials in which 4 main aspects (with many items each) are considered: (1) basic descriptive material; (2) the study protocol; (3) the analysis of the data; (4) clear presentation of the results.² They considered randomization and blinding as important criteria. The criteria they used are very similar to the ones used in this thesis. Chalmers et al. used aspects 2, 3, and 4 to create a quantitative score with relative weights of 0.6, 0.3 and 0.1 respectively. In our reviews of homoeopathy and ginkgo the corresponding weights used were 0.85 for the study protocol and 0.15 for the analysis and the presentation of the results. A minor difference between the scoring systems is that we also gave points when endpoints used were sensible (i.e. clinically relevant). Chalmers et al. present an example of a trial with an overall score of 0.84, which in their experience is a good score. They mention that they have some trials with scores in the 0.90's. Emerson et al. sorted 107 randomized clinical trials according to their quality scores (using the criteria of Chalmers et al.)² and formed three groups: "lower" (scores of 0-0.39), "middle" (0.40-0.51) and "higher" (0.52-0.89).³ Examples of the evaluated trials are antibiotic treatment in gastro-intestinal surgery, dipyridamole for angina pectoris, antiplatelets for prevention of stroke, progesterone in pregnancy and others. They found that the level of quality increased at a rate of 9% per decade for three decades (starting 1960), averaging 0.51 for the 1980's. DerSimonian et al. used 11 items which are similar to Chalmers et al's and investigated whether these items were mentioned in clinical trials published in four medical journals (N Engl J Med, JAMA, Lancet, Br Med J).⁴ In 67 clinical trials (published in 1979-1980), only 56% of all 11 items were clearly reported, 10% were ambiguously mentioned and 34% were not reported at all. These examples are comparable to the scores of the quality index used in our reviews. Two differences are our emphasis on the numbers of participants included (the other authors demand power calculations) and side-effects (which we comment on in the results sections of our reviews).

We never performed statistical pooling of trial results. In all subject matters covered there were too many differences between trials as for the characteristics of the participants, the interventions, and the effect measurements. In any case, pooling can only be considered for well performed trials; it makes no sense to pool evidence of high and low quality. Furthermore, if there is serious suspicion of publication bias, pooling of a selection of trials with positive results makes little sense.

Publication bias

For our collection of controlled trials, we had the opportunity to visit researchers and pharmaceutical companies in The Netherlands, the U.K., France, Germany and Switzerland. We visited 6 manufacturers of food supplements, 4 institutions where research into food supplements is performed, and 8 leading researchers. In almost all instances these visits led to the identification of controlled trials that were new to us. Correspondence

and telephone calls are simply not sufficient if one is trying to identify all relevant trials. Also, to ask personally whether there have been any unpublished trials, often leads to the identification of new trials. This is a direct confirmation of the suspicion that trials exist which remain unpublished. A frequently encountered reason for non-publication was that it was considered logical that no positive results were found because, in hindsight, the wrong patients, the wrong intervention, or the wrong effect measurements had been used. Another interesting fact is that in all cases the newly identified trials did not involve large numbers of patients.

Biological plausibility of the mechanisms of action

There are numerous examples in medicine in which innovations were delayed because there was no plausible mechanism of action. Nowadays, clinical trials are used to assess the efficacy of interventions, and generally the results are respected and implemented in practice even if there is no explanation why a treatment would work. Before reading an article about efficacy research, a reader has some prior opinion about the efficacy of the intervention reported on. This may be based on many factors: what was taught at the university, knowledge about the proposed mechanisms of action (biological plausibility), personal experience in treating patients, and also the results of earlier trials.

When additional evidence is presented, be it an article about a clinical trial or a review, this new information can change the belief in the efficacy.⁵ Biological plausibility may play a more important role in the magnitude of the change in belief than expected. In the case of homoeopathy, there was an unexpectedly large body of evidence indicating that it might be efficacious for at least some indications. However, our prior belief was extremely low, because often homoeopathic medicines do not contain a single active molecule, and therefore we still did not seriously believe in the efficacy of homoeopathy, in spite of positive results from clinical trials. The question arises whether in some cases the results of clinical trials are only to be believed if there is some notion of a mechanism of action? The results of the review of homoeopathy would probably be interpreted differently, if laboratory studies would show convincing evidence that there is some action of high potencies. A conclusion might be that in this case biological plausibility is more important than the results of clinical trials. The future may learn whether this point of view is justified.

References

1. Knipschild P, Leffers P, Feinstein AR. The qualification period. *J Clin Epidemiol* 1991;44:461-4.
2. Chalmers TC, Smith Jr H, Blackburn B, et al. A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials* 1981;2:31-49.
3. Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Controlled Clinical Trials* 1990;11:339-52.
4. DerSimonian R, Charette LJ, McPeck B, Mosteller F. Reporting on methods in clinical trials. *N Engl J Med* 1982;306:1332-7.
5. Knipschild P, Kleijnen J, Riet ter G. Zur Glaubwürdigkeit alternativer Medizin. *Skeptiker* 1990;3(3):4-8.

SUMMARY

Many statements about the efficacy of food supplements exist. For example, there are claims that garlic lowers cholesterol levels; vitamin C prevents colds; vitamin B6 reduces the symptoms of the premenstrual syndrome; vitamins improve the IQ; and homoeopathy reduces the symptoms of hayfever. Is there scientific evidence supporting these claims?

In this dissertation the scientific evidence on eight of such claims has been assessed. We did not personally perform any experiments, but instead chose to review the evidence that can be found in the medical literature. The most important criterion for articles to be included in the reviews was that the effects in the index group should be compared with those in a control group (mostly a placebo group) in order to assess objectively whether the use of food supplements makes any sense.

Making such assessments, the first step we took was the collection of all relevant articles. This was achieved by searching computerized databases containing all articles from major medical journals which have been published in the last 25 years. Subsequently, references in the identified articles were checked to find additional publications. However, we presumed that more articles exist, published in obscurer journals, books or dissertations. In order to obtain a complete overview of all relevant information, we tried to retrieve articles from the latter set as well. To obtain this information, we corresponded with leading researchers and major pharmaceutical companies in the respective fields of interest. We personally visited several of these researches and companies, and also some specialized libraries, in an attempt to find even the last single publication. Chapter 2 describes the procedure and the yield of a computersearch. Subsequently, the number of articles found in a computersearch was compared with that found with more extensive search methods. It turned out that with computersearches only a minority of all relevant articles can be identified. However, if the references in the identified articles are followed-up, an acceptable image of the state-of-affairs is obtained: most of the articles with a higher quality will be identified. If, however, a complete review is necessary, more elaborate search methods have to be used. Of course, this conclusion can only be held up for the subjects covered in this dissertation.

Chapter 3 reports on a survey among some 300 Dutch GP's. They were asked to state their belief in the efficacy of food supplements and many alternative procedures. They could give their opinion on a scale from 0 to 10. Almost half of them believed that homoeopathy can be efficacious in the treatment of upper respiratory infections. In general the GP's did not believe that food supplements are efficacious; only a minority thought that vitamin B6 is useful for premenstrual complaints (36%) and that vitamin C is helpful in colds (20%).

Unfortunately, articles published in medical journals differ in quality to a great extent. Of course, only high quality evidence can be used for the assessment of the efficacy. A list of criteria of good methodology was used for the evaluation of the quality of the experiments. The state-of-affairs on the efficacy of several food supplements will be presented in the chapters 4-12.

Chapter 4 provides an overview of all trials on the prevention and treatment of colds by vitamin C. No less than 61 studies have been published. From these trials it appears that vitamin C is not effective in the prevention of colds. However, there may be a small positive effect on duration and severity of symptoms. Possibly, the duration of a cold that would normally last five days might be shortened by half a day. An effect of similar size might be expected on the severity of the symptoms.

Chapter 5 presents trials on the influence of vitamin E on angina pectoris and intermittent claudication. For angina pectoris, six controlled trials have been performed, but no positive effects were reported. On the other hand, 4 out of 5 trials in patients with intermittent claudication reported positive results. However, in these trials, performed in the 1950's and 1960's, a limited number of participants was used. Consequently, it would be interesting to see the results of new, larger, trials.

In chapter 6, research into the effects of niacin and vitamin B6 on mental functioning is described. Many categories of patients have been used during the years, varying from effects in schizophrenics to those in autistic children. Also trials about the effects on IQ in healthy schoolchildren have been performed. The general picture is not encouraging. In most trials, no differences compared with placebo treatment were found. The only indication in which further trials seem warranted is autism: three small trials reported positive results of a high dose of vitamins combined with magnesium supplementation. The influence of vitamin B6 on the premenstrual syndrome is evaluated in chapter 7. We found 12 controlled trials assessing the effect on complaints such as abdominal bloating, breast tenderness, headache, edema, fatigue and depression. The results are contradictory: in 3 studies positive effects were reported, although the symptoms had not disappeared completely; in 5 studies some positive effects were found, but the overall result was disappointing; and in 4 studies no positive effects at all were found.

Chapter 8 reports on trials evaluating the effect of evening primrose oil on many diseases such as multiple sclerosis, rheumatoid arthritis, premenstrual syndrome and atopic eczema. There is no substantial evidence of the efficacy of evening primrose oil in any of these indications. In all trials, several methodological shortcomings exist, but for some indications such as the premenstrual syndrome it would be interesting to see the results of new, better studies.

The effects of fresh garlic and garlic supplementation are evaluated in chapters 9 and 10. Chapter 9 provides an overview of all trials on fresh garlic and of the earlier trials with garlic supplements. Almost all trials on fresh garlic show positive effects on cardiovascular risk factors: a lowering of cholesterol levels, decreased platelet aggregation and increased fibrinolytic activity. These effects are found using dosages equivalent to 7-28 cloves a day. It would be very interesting to see the results of trials with lower dosages. Also the effects of supplements, without the garlic odor, would be interesting. Most of the trials presented in chapter 10 have only recently been published, and the results confirm that the positive effects of fresh garlic are also found with at least some supplements. However, before claiming that cardiovascular disease itself is influenced, trials with morbidity and even mortality outcome measurements have to be performed.

Chapter 11 provides a review of studies with extracts of the leaf of the ginkgo biloba tree. Ginkgo is very popular in Germany and France. It is used for symptoms of elderly people such as difficulties with concentration and memory, tinnitus and dizziness. Ginkgo is also prescribed in intermittent claudication. Many controlled trials have been performed, and especially in cerebral insufficiency there is much evidence in favour of positive effects. However, further trials must be performed for a more detailed assessment of the efficacy.

Controlled trials on the efficacy of homoeopathy are reviewed in chapter 12. There are more than 100 controlled trials on the efficacy of homoeopathy. In most of these studies homoeopathic treatment has been compared with placebo treated control groups.

Amongst others, it concerns patients with upper respiratory tract diseases, gastro-intestinal complaints and patients with pain complaints, for instance from rheumatic diseases. Almost 80% of these trials report positive results, irrespective of the applied variety of homoeopathy. Classical homoeopathy, where every patient receives individual therapy, was investigated in 14 trials. The other studies concerned modern varieties of homoeo-

pathy, in which all patients receive the same, often highly diluted, treatment. Homoeopathic doctors call this a potency, because the diluting process is accompanied by vigorous shaking, in order to increase rather than decrease the effectiveness. This is contrary to what regular doctors would expect to happen. The quality of the studies is generally disappointing, for instance because in many trials only small numbers of patients participated. However, there are also some 23 studies which are of at least reasonable, but in some cases of very high quality. In 15 of these 23 studies, a positive effect of homoeopathic treatment was found compared with the control group. It may be possible that they present a biased picture of the true state-of-affairs, because trials with positive results might have had a greater chance of being published. Further trials are necessary, but they only make sense if they are of high quality, with large numbers of participants.

In chapter 13 publication bias is discussed. Publication bias occurs if the results from studies which have not been published, are different from the published ones. From a Bayesian viewpoint, it also concerns non-publication of studies with similar results as the published ones, because the strength of the evidence will be influenced. Publication bias complicates the interpretation of reviews and meta-analyses. If favourable results are published more often, there will be an overestimation of the effects of a treatment. This this serious problem is discussed in detail.

Chapter 14, finally, describes some general problems and ideas about criteria-based meta-analysis, and we compare our criteria with the criteria used by other research groups. Furthermore we elaborate on some personal experiences with publication bias and discuss the results of clinical trials with respect to the biological plausibility of the mechanisms of action.

It can be concluded that, generally, the evidence on the efficacy of food supplements is disappointing. Positive results have been reported for garlic, ginkgo and homoeopathy, but also for these interventions further trials are necessary. For modern reviews of the literature similar standards must be applied as for reports of controlled trials or other scientific research.

SAMENVATTING

Regelmatig worden uitspraken gehoord over voedingssupplementen. Voorbeelden zijn: "knoflook reinigt het bloed"; "vitamine C helpt tegen verkoudheid"; "vitamine B6 is goed voor menstruatie klachten"; "vitamines verhogen het IQ"; en "homeopathie helpt bij hooikoorts". Maar bestaan er ook bewijzen uit wetenschappelijk onderzoek dat dit inderdaad het geval is?

In dit proefschrift worden voor acht van deze claims de bewijzen voor de werkzaamheid op een rijtje gezet. Hierbij gaat het niet om eigen experimenten bij patiënten, maar wel om een beoordeling van alle onderzoek dat in de medische literatuur beschreven is. De belangrijkste eis waaraan dergelijk onderzoek moet voldoen, is de volgende: er dient een vergelijking gemaakt te zijn tussen een groep patiënten die het te onderzoeken middel kreeg en een soortgelijke groep die in de meeste gevallen een nepmiddel (placebo) kreeg. Op deze manier kan geprobeerd worden objectief vast te stellen of het nemen van dergelijke voedingssupplementen zin heeft of niet.

Om het e.e.a. te kunnen beoordelen moeten allereerst alle relevante onderzoeken worden opgespoord. Dit kan met hulp van computerbestanden, die alle artikelen bevatten uit vooraanstaande medische tijdschriften die in de afgelopen 25 jaar gepubliceerd zijn. Vervolgens kunnen andere artikelen van soortgelijk onderzoek, waarnaar in deze artikelen verwezen wordt, worden opgevraagd. Maar misschien is er nog meer interessante informatie, gepubliceerd in tijdschriften, boeken of proefschriften die niet zo bekend zijn. Wij hebben geprobeerd ook deze laatste categorie onderzoeken op te sporen, om het plaatje zo compleet mogelijk te maken. Hiervoor is gecorrespondeerd met bv. vooraanstaande onderzoekers op de betreffende vakgebieden en belangrijke farmaceutische firma's. Ook hebben wij enkele van deze onderzoekers en firma's bezocht, en zijn we naar enkele gespecialiseerde bibliotheken geweest om ook het laatste relevante onderzoek boven water te halen. Hoofdstuk 2 beschrijft hoe een computersearch in zijn werk gaat en wat de opbrengst is. Vervolgens hebben we de opbrengst van zo'n computersearch vergeleken met het aantal artikelen dat gevonden wordt na uitgebreidere speuracties. Het blijkt dat met een computersearch slechts een minderheid van alle onderzoeken wordt opgespoord. Indien echter alle citaties in de gevonden artikelen gecontroleerd worden, ontstaat een redelijk beeld van de stand van zaken: de meeste van de betere onderzoeken worden op deze manier wel gevonden. Wil men echter een compleet beeld van de stand van zaken krijgen, dan zal men aanvullende speuracties moeten ondernemen. Uiteraard kan deze conclusie alleen gelden voor de onderwerpen die in dit proefschrift beschreven worden.

In hoofdstuk 3 wordt verslag gedaan van een enquête onder zo'n 300 Nederlandse huisartsen. Aan hun werd gevraagd om met een rapportcijfer aan te geven in hoeverre ze in de werkzaamheid van o.a. voedingssupplementen geloofden. Bijna de helft ziet wel iets in homeopathie bij de behandeling van bovenste-luchtweginfecties. Huisartsen geloven in het algemeen niet in de werkzaamheid van vele voedingssupplementen; alleen vitamine B6 bij premenstruele klachten (36%) en vitamine C bij verkoudheid (20%) krijgen van een minderheid een voldoende.

Helaas is het niet zo dat alle onderzoek dat in medische tijdschriften gepubliceerd is, ook kwalitatief goed is uitgevoerd. Integendeel, er blijken grote verschillen in de kwaliteit van onderzoek te bestaan, en het spreekt voor zich dat alleen goed uitgevoerd onderzoek nuttige informatie levert. In de volgende hoofdstukken wordt de stand van zaken over de werkzaamheid van een aantal voedingssupplementen gepresenteerd. Hierbij zijn de gevonden onderzoeken aan de hand van een aantal criteria op hun kwaliteit beoordeeld.

Hoofdstuk 4 beschrijft het onderzoek naar de werkzaamheid van vitamine C bij het ontstaan en de behandeling van verkoudheden. Er blijken maar liefst 61 studies te zijn verricht. De uitkomsten laten zien dat het geen nut heeft vitamine C te nemen om verkoudheden te voorkómen. Als er al van enig positief effect sprake is, dan is dat bij de behandeling van verkoudheden. Mogelijk bekort vitamine C de duur van een verkoudheid die normaal 5 dagen zou duren, met een halve dag. Voor de symptomen zou een vermindering in dezelfde orde van grootte verwacht kunnen worden.

Hoofdstuk 5 gaat over vitamine E bij hart-vaatziekten. We vonden 6 studies met controlegroepen voor de invloed op klachten van pijn op de borst (angina pectoris). De uitkomst was dat geen positieve effecten te verwachten zijn. Voor etalageziekte (claudicatio intermittens) lieten 4 van de 5 studies positieve resultaten zien. Omdat het echter vooral kleine studies betreft uit de jaren vijftig en zestig, zou e.e.a. nog eens over moeten worden gedaan met grotere groepen patiënten.

Hoofdstuk 6 beschrijft onderzoek naar de invloed van vitamine B3 en B6, al dan niet in combinatie met andere vitamines en mineralen, op het geestelijk functioneren. Dit is bij vele categorieën patiënten getest, variërend van patiënten met schizofrenie tot autistische kinderen. Ook zijn de effecten op het IQ van normale, schoolgaande kinderen getest. Het algemene beeld is niet positief. Er blijken geen verschillen op te treden tussen patiënten die met vitamines en patiënten die met een placebo worden behandeld. Het enige lichtpuntje is dat bij 3 studies met autistische kinderen enige gunstige effecten gevonden zijn. De betreffende studies zijn echter van matige kwaliteit, en de uitkomst zal dus in verder onderzoek bevestigd of ontkend moeten worden.

Hoofdstuk 7 gaat over vitamine B6 bij vrouwen met een premenstrueel syndroom. Dit zijn klachten die voor de menstruatie kunnen optreden, zoals buikpijn, een gespannen gevoel in de borsten, vochtophoping en snel geïrriteerd zijn. Er blijken 12 onderzoekingen naar de effecten van vitamine B6 bij deze klachten te bestaan. De uitkomsten zijn nogal tegenstrijdig: 3 onderzoeken melden een positieve invloed op de klachten, maar de vrouwen werden niet helemaal beter. Vijf andere onderzoekingen laten voor enkele symptomen wel wat positieve effecten zien, maar eigenlijk is het resultaat teleurstellend en tenslotte zijn er ook 4 onderzoekingen waarin helemaal geen positieve effecten gevonden worden.

Hoofdstuk 8 bespreekt de invloed van teunisbloemolie bij allerlei aandoeningen zoals multipale sclerose, reumatoïde arthritis, premenstrueel syndroom en eczeem. Geen enkele claim blijkt voldoende te zijn aangetoond. Ook blijkt dat het een en ander valt aan te merken op de kwaliteit van het meeste onderzoek. Voor een aantal aandoeningen zoals premenstrueel syndroom zou het interessant zijn om nieuw, beter onderzoek te verrichten.

In de hoofdstukken 9 en 10 worden onderzoekingen besproken naar de effecten van knoflook en knoflookpillen op risicofactoren voor het ontstaan van hart-vaatziekten. Bijna alle onderzoek met verse knoflook laat positieve resultaten zien. Zo wordt het cholesterolgehalte verlaagd, en wordt de bloedstolling gunstig beïnvloed. Daarbij gaat het wel om doseringen van 7 tot 28 teentjes per dag! Het zou dus interessant zijn om de effecten van knoflookpillen, die wat minder geuren, te onderzoeken. Helaas is het onderzoek dat voorhanden is over deze pillen, nog niet van voldoende kwaliteit om definitieve uitspraken over de werkzaamheid te kunnen doen. Bovendien moet er rekening mee gehouden worden dat beïnvloeding van risicofactoren niet altijd wil zeggen dat het optreden van hart-vaatziekten zelf ook beïnvloed wordt. Dit zal in apart onderzoek vastgesteld moeten worden.

Hoofdstuk 11 gaat over studies met extracten van de bladeren van de ginkgo boom. Dit middel, dat in Duitsland en Frankrijk zeer populair is, wordt toegepast bij vele verschijnselen die nogal eens bij oudere mensen voorkomen zoals o.a. vergeetachtigheid, moeilijk-

heden met concentreren, oorsuizen en duizeligheid. Ook wordt het toegepast bij etalageziekte (claudicatio intermittens), welke gekenmerkt wordt door pijn in de benen bij het lopen. Er is al veel onderzoek naar de effecten van dit middel gedaan, en het lijkt erop dat inderdaad positieve effecten verwacht kunnen worden. Enkele goed uitgevoerde, nieuwe onderzoeken zouden e.e.a. definitief kunnen aantonen.

Hoofdstuk 12 bespreekt de effectiviteit van homeopathie bij allerlei aandoeningen. Er blijken meer dan 100 onderzoeken te bestaan over de werkzaamheid van homeopathie. Daarbij gaat het om studies bij mensen waarbij de resultaten van homeopathie vergeleken zijn met een nep-behandeling (placebo) of andere behandelingen. Het merendeel is de laatste 10 jaar uitgevoerd, vooral in Groot-Brittannië, Duitsland en Frankrijk. Het betreft patiënten met onder meer ziekten van de luchtwegen, van de darmen en met pijnklachten, waaronder rheuma.

Bijna 80% van deze onderzoeken laten positieve resultaten zien, ongeacht de gekozen vorm van homeopathie. Klassieke homeopathie, waarbij iedere patient zijn eigen middel krijgt, is in 14 studies onderzocht. De overige studies betreffen veel toegepaste moderne varianten waarbij groepen patiënten met dezelfde, soms extreme verdunning behandeld worden. Homeopathen spreken in dat verband over potentiëring, omdat het proces van verdunnen vergezeld gaat met schudden, zodat de werkzaamheid eerder zou toenemen dan afnemen. Daarbij zetten wij grote vraagtekens.

De kwaliteit van de studies is in het algemeen laag, onder meer omdat te weinig patiënten aan zo'n onderzoek deelnamen. Niettemin zijn er 23 studies van redelijke tot zeer goede kwaliteit. In deze groep van beter uitgevoerde studies komt homeopathie 15 keer als gunstig te voorschijn. Deze zijn trouwens vaak niet gepubliceerd in bekende medische tijdschriften. Mogelijk is er een vertekening opgetreden doordat voor homeopathie gunstige resultaten een grotere kans maakten om gepubliceerd te worden.

Gezien de grote populariteit van homeopathie is nieuw onderzoek dringend noodzakelijk. Dergelijk onderzoek heeft echter alleen zin als het nagenoeg perfect wordt uitgevoerd bij grote aantallen patiënten.

Hoofdstuk 13 gaat in op het probleem van publicatiebias. Het zou zo kunnen zijn dat van alle onderzoeken over een bepaald onderwerp slechts een deel wordt gepubliceerd, en de rest in een bureaula verdwijnt. Stel nu dat de uitkomst van het onderzoek de kans beïnvloedt dat het gepubliceerd wordt of niet (bijvoorbeeld indien studies waar niets uitkomt, vaker in een la verdwijnen). In dat geval geven overzichten van de gepubliceerde onderzoeken een vertekend beeld van de werkelijkheid. In dit hoofdstuk wordt dit vermoedelijk ernstige probleem uitvoerig belicht.

In hoofdstuk 14 worden tenslotte nog enkele zaken besproken, die bij het literatuuronderzoek naar voren kwamen. Daarbij gaat het om problemen bij en nadelen van onze kwalitatieve beoordeling van de onderzoeken. Ook bespreken wij enkele persoonlijke ervaringen met publicatiebias. Tot slot wordt ingegaan op de invloed van kennis van werkingsmechanismen (hoe een middel het effect veroorzaakt) bij de beoordeling van het bewijs van werkzaamheid.

Algemeen kan gesteld worden dat de bewijzen voor de werkzaamheid van vele voedings-supplementen teleurstellend zijn. Voor knoflook, ginkgo en homeopathie lijken de uitkomsten gunstig, maar ook voor deze supplementen is verder onderzoek noodzakelijk. Moderne overzichtsartikelen van de literatuur dienen aan dezelfde kwaliteitsnormen te voldoen als effectiviteitsonderzoek. Dit betekent een duidelijke omschrijving van de opsporingsmethoden van de onderzoeken, en aan welke criteria deze moeten voldoen, hoe hun kwaliteit is beoordeeld, en hoe men tot conclusies komt.

1. $\frac{1}{x^2} = x^{-2}$
2. $\frac{d}{dx} x^{-2} = -2x^{-3}$
3. $= -2x^{-3}$
4. $= -\frac{2}{x^3}$
5. $= -\frac{2}{x^3}$
6. $= -\frac{2}{x^3}$
7. $= -\frac{2}{x^3}$
8. $= -\frac{2}{x^3}$
9. $= -\frac{2}{x^3}$
10. $= -\frac{2}{x^3}$

CURRICULUM VITAE

Jos Kleijnen werd op 23 februari 1963 te Heerlen geboren. In 1981 werd aan het Dr. Moller College te Waalwijk het Atheneum B diploma behaald. Vervolgens studeerde hij geneeskunde aan de Rijksuniversiteit Limburg, alwaar hij in 1987 het artsexamen aflegde. Sindsdien is hij als toegevoegd onderzoeker werkzaam bij de vakgroep Epidemiologie/Gezondheidszorgonderzoek van de Rijksuniversiteit Limburg. In het kader van het door het Ministerie van WVC gesubsidieerde project "Literatuurstudies effectiviteit alternatieve geneeswijzen" schrijft hij overzichtsartikelen over de effectiviteit van verschillende alternatieve geneeswijzen en voedingssupplementen. Zijn postacademische scholing bestond ondermeer uit deelname aan cursussen op het gebied van de epidemiologie. Vanaf 1987 heeft hij, naast de in dit proefschrift opgenomen artikelen, gepubliceerd over acupunctuur, meta-analyse als reviewmethode en over placebo-effecten.

DANKWOORD

Mijn dank gaat uit naar Paul Knipschild en Gerben ter Riet.
Als zij er niet waren, was dit proefschrift er niet.