Key Issues in Clinical and Epidemiological Research in Complementary and Alternative Medicine – a Systematic Literature Review

H. Felix Fischer a  Florian Junne b  Claudia Witt a  Klaus von Ammon c  Francesco Cardini d  Vinjar Fønnebø e  Helle Johannessen f  George Lewith g  Bernhard Uehleke h,i  Wolfgang Weidenhammer j  Benno Brinkhaus a

a Institute for Social Medicine, Epidemiology and Health Economics, Charité University Medical Center, Berlin, Germany
b Department for Psychosomatic Medicine and Psychotherapy, University Hospital Tübingen, Germany
c Institute of Complementary Medicine KIKOM, University of Bern, Switzerland
d Health and Social Agency of Emilia Romagna Region, Bologna, Italy
e National Research Center in Complementary and Alternative Medicine, (NAFKAM), Department of Community Medicine, University of Tromsø, Norway
f Institute of Public Health, University of Southern Denmark, Odense, Denmark
g Complementary and Integrated Medicine Research Unit, University of Southampton, UK
h Institute of Complementary Medicine, University Hospital Zurich, Switzerland
i University of Health and Sports, Berlin,
j Competence Centre for Complementary Medicine and Naturopathy, Klinikum rechts der Isar, Technical University München, Germany

Keywords
Clinical research · Epidemiological research · Research methodology · Complementary and alternative medicine · Systematic review

Summary
Background: In the last 2 decades there has been a large increase in publications on complementary and alternative medicine (CAM). However, CAM research methodology was heterogeneous and often of low quality. The aim of this systematic review was to investigate scientific publications with regards to general issues, concepts and strategies. We also looked at research priorities and methods employed to evaluate the clinical and epidemiological research of CAM in the past to identify the basis for consensus-based research strategies. Methods: We performed a systematic literature search for papers published between 1990 and 2010 in 7 electronic databases (Medline, Web of Science, PsychArticles, PsychInfo, CINAHL, EMBASE and Cochrane Library) on December 16 and 17, 2010. In addition, experts were asked to nominate relevant papers. Inclusion criteria were publications dealing with research methodology, priorities or complexities in the scientific evaluation of CAM. All references were assessed in a multistage process to identify relevant papers. Results: From the 3,279 references derived from the search and 98 references contributed by CAM experts, 170 papers fulfilled the criteria and were included in the analysis. The following key issues were identified: difficulties in past CAM research (e.g., randomisation, blinding), utility of quantitative and qualitative research methods in CAM, priority setting in CAM research and specific issues regarding various CAM modalities. Conclusions: Most authors vote for the use of commonly accepted research methods to evaluate CAM. There was broad consensus that a mixed methods approach is the most suitable for gathering conclusive knowledge about CAM.

Introduction
CAMbrella is a European Union (EU)-funded coordinated action in the field of complementary and alternative medicine (CAM). To address the increasing use of CAM and the lack of scientific knowledge concerning CAM use, the CAMbrella Work Package 7 (WP7) group is developing a ‘roadmap for further clinical and epidemiological research in CAM’. Here
we report the results of a systematic literature review on general issues in CAM research as a first step towards the development of a research roadmap. Research in CAM has been a controversial topic (for a broad overview on CAM research see [1]), and our aim was to create a comprehensive evaluation and analysis of the methodological and conceptual issues involved.

We therefore performed a systematic review of literature dealing with the complexities and general methodological issues involved in the evaluation of CAM in clinical and epidemiological research. Ultimately, the outcome of this review, the subsequent discussion and the final roadmap for further research, should lead to a basis and framework for further CAM research in Europe.

Methods

A structured systematic literature review was conducted. Before starting this review, a systematic review protocol was developed (initial draft by the WP7 leader), which was submitted to the whole WP7 group for notes and suggestions for changes. The final version of the review protocol (including the search terms) was approved by the whole WP7 group.

Literature Search

In 2010, 7 electronic databases (Medline, Web of Science, PsychArticles, PsycInfo, CINAHL, EMBASE and Cochrane Library) were searched for relevant articles published between 1990 and 2010 (until December 16/17). Table 1 shows the search terms entered into the databases. In addition to the database search, all experts and the advisory board involved in the CAMbrella project were asked to submit any relevant publications.

Selection

Duplicates were excluded. We mainly aimed for full articles and original works, but comments, editorials, letters and ‘grey’ literature were included when a substantial original contribution to the topic was found. The title and abstract of the remaining references were screened by 1 researcher (Florian Junne) to exclude irrelevant references that were not related to CAM at all, not in a European language, on basic research only or on animal studies. Secondly, the title and abstract of the remaining articles were evaluated by 2 reviewers (Florian Junne and Felix Fischer) to identify publications that included investigations, analysis, discussion, proposals or statements concerning the following: i) qualitative and quantitative methods, ii) clinical and epidemiological research methodology, iii) priorities or priority setting or iv) methodological complexities involved in the scientific evaluation of CAM.

Articles with a corresponding judgment from both reviewers were included in further analysis. Kappa as measure of inter-rater agreement was calculated. For non-corresponding judgments, the 2 reviewers discussed the title and abstract of the publication until an agreement regarding inclusion or exclusion was achieved. Publications contributed by CAM experts (additional references were contributed by the authors of the review and the CAMbrella Advisory Board members Nora Laubstein, Ton Nicolai, Peter Zimmermann and Stephen Gordon) were also reviewed and included in full-text analysis if they met the inclusion criteria after rating of title and abstract.

Full-Text Analysis and Data Extraction

All included publications entered full-text analysis. The eligibility of the publications was re-examined with respect to the above-mentioned inclusion/exclusion criteria. At this stage, publications were also excluded if 1 of the following additional exclusion criteria was fulfilled: (i) it mainly addressed research methodology of basic and experimental research; (ii) it primarily addressed the reporting of clinical trials; (iii) it primarily assessed methodological quality/riour of CAM-evaluation trials; (iv) it presented a case study or abstract only; or (v) it mainly reported a specific study design or research tool.

Results

The literature search resulted in 3,279 hits and CAMbrella members contributed 98 additional references. After the exclusion process, 170 studies were included in the qualitative synthesis. See figure 1 for additional information.
Practical Problems in Research into CAM

We found a large number of publications dealing with practical problems when conducting research in CAM. These problems and relevant references are categorised in detail in table 2.

Choice of Research Methods

The choice of research method depends on the question asked [2–6]. In some publications, explicit research questions and appropriate methods were given by the authors [2, 5, 7–11]. However, there was clear agreement about the value of different research methods in CAM research [12–16]. Most authors suggested a research-question-driven integration of diverse methods into the research agenda [6–8, 13, 16–29].

Quantitative Research Methods

Randomised controlled trials (RCTs) are considered to be the gold standard to assess specific effects and efficacy and to determine causal relationships in biomedical research. Most authors stated that RCTs with high methodological quality are possible in the field of CAM and can therefore produce valid data [4, 11, 17, 30–48], but they must be rigorously performed and CAM-specific challenges must be addressed, such as the lack of external validity due to strict standardisation of diverse treatments and study participants [3, 30, 32, 37, 38, 41–43, 48–59]. However, a consensus emerged that clearly implied that RCTs do not answer all research questions [28, 36, 52, 60–64] and are expensive to conduct [18, 30, 32, 52, 65, 66]. Some authors argue that placebo-controlled RCTs might be inappropriate for some specific CAM modalities [67–69]: a position that has raised considerable controversy [47, 70]. Integration of diverse research methods [2, 12, 27, 38, 61, 71], preference trials [3, 72, 73] or the use of different outcome measures [74, 75] could help overcome these shortcomings. Feasibility studies are a vital preliminary phase in the design of high-quality RCTs with adequate power [3, 4, 10, 76–79]. When individualised and standardised treatments are to be compared [3], or if specific and non-specific effects need to be separated [80], RCTs can be extended to more than 2 treatment arms to account for preference towards a specific treatment in preference trials [20, 28, 72, 73].

Pragmatic trials – as promoted in Comparative Effectiveness Research (CER) – can be conducted to assess outcomes...
### Table 2. Problems experienced in CAM research

<table>
<thead>
<tr>
<th>Problem</th>
<th>Short description</th>
<th>Possible solutions</th>
<th>Relevant references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity of interventions</td>
<td>therapies in CAM typically consisting of a number of different procedures and/or interventions; isolation of parts may lead to underestimation of effect</td>
<td>research into overall effects/effectiveness; refined methodological approaches</td>
<td>[12, 17, 30, 31, 49, 60, 87, 99, 130, 146, 149]</td>
</tr>
<tr>
<td>Assessment of specific and non-specific effects</td>
<td>unclear nature of unpecific effects in CAM, little external validity of research into specific effects so far, RCTs rule out possibly important non-specific effects</td>
<td>development of a clear definition of nonspecific effects, consideration of specific/non-specific effects in trial design, prioritisation of effectiveness research</td>
<td>[2, 12, 17–19, 31–33, 30, 61, 62, 72, 80, 130, 136, 137, 151, 165]</td>
</tr>
<tr>
<td>Choice of control group(s)</td>
<td>appropriate control group depends on the research question asked (especially placebo); specific control group conditions not feasible in some cases</td>
<td>different control conditions are possible, even within the same trial. Placebo treatments must be carefully developed</td>
<td>[18, 30, 34–37, 51–53, 63, 72, 99, 101, 143, 144, 151, 152, 158, 166, 167]</td>
</tr>
<tr>
<td>Randomisation</td>
<td>randomisation is desirable, but might be hindered by patients' preferences making them unwilling to participate in studies when allocated to placebo treatment</td>
<td>take preferences into account in trial design (placebo trial) and statistical analysis; when randomisation is impossible, assess baseline differences</td>
<td>[20, 30, 32, 33, 37–39, 53–55, 62, 67, 72, 84, 87, 88, 111, 115, 127, 146, 147, 168]</td>
</tr>
<tr>
<td>(Double-)blinding</td>
<td>(double-)blinding is desirable, but not achievable in all trials of CAM treatments (e.g. blinding of an acupuncturist)</td>
<td>assess success of blinding technique; if impossible, blind outcome assessor, data analyst, diagnostican</td>
<td>[2, 3, 33, 35, 37, 40–43, 49, 51–53, 55, 68, 101, 102, 115, 152, 154, 169]</td>
</tr>
<tr>
<td>Handling of different diagnostic frameworks</td>
<td>treatment allocation within studies could differ between different diagnostic systems of CAM and conventional medicine</td>
<td>differences between diagnostic systems must be assessed and should be taken into account when allocating treatment</td>
<td>[33–35, 43, 44, 49, 103, 123, 134, 170]</td>
</tr>
<tr>
<td>Definition of treatment; standardisation vs. individualised treatment</td>
<td>standardisation leads to loss of external validity, since most CAM treatments are considered as necessarily individualised; individualised treatment hinders reproducibility of trials</td>
<td>use of semi-standardised treatment regimens; implementation of individualised and standardised treatments as study arms</td>
<td>[3, 13, 17, 18, 21, 31, 34, 39, 42, 49, 68, 134, 152, 153, 166, 171, 172]</td>
</tr>
<tr>
<td>Time frame for expected results</td>
<td>observation time for CAM studies might need to be longer, e.g. in the treatment of chronic illness</td>
<td>extending study duration and regular follow-ups</td>
<td>[17, 30, 35, 39, 49, 65, 88]</td>
</tr>
<tr>
<td>Choice of outcome parameters</td>
<td>treatment effects/outcomes in CAM might be different from conventional medicine</td>
<td>objective and subjective outcomes should be assessed (if possible) and cover different domains; when outcomes are unclear different potential outcomes should be considered</td>
<td>[20–23, 30, 39, 41, 44, 49, 56, 67, 76, 89, 97, 101, 107, 130, 173, 174]</td>
</tr>
<tr>
<td>Study setting and treatment providers</td>
<td>CAM is often applied by practitioners with little experience in research; treatment provided by different practitioners might be hard to standardise</td>
<td>research needs to be conducted in collaboration with experienced researchers and clinicians</td>
<td>[14, 74, 86, 89, 138]</td>
</tr>
<tr>
<td>Lack of knowledge underlying the mechanism of interventions</td>
<td>lacking theoretical basis of treatments complicates the planning of valid studies and might compromise results</td>
<td>implementation of research into foundations of CAM</td>
<td>[4, 13, 17, 24, 55, 57, 58, 76–78, 93, 108, 152, 156, 160, 175]</td>
</tr>
<tr>
<td>Inconclusive study results</td>
<td>study results are controversial, e.g. in homeopathy, acupuncture &amp; dietary supplements</td>
<td>development of guidelines in trial design, enhance methodological quality by development of research infrastructure</td>
<td>[5, 7, 15, 43, 45, 52, 134, 144, 156, 159]</td>
</tr>
</tbody>
</table>

RCT = Randomised controlled trial.
of a treatment within a real world clinical setting (clinical effectiveness) [36, 69, 81–83]. Pragmatic trials enable comparison of clinical treatment alternatives, inclusion of a wide variety of patients in diverse practice settings and address a broad range of patient relevant outcomes [2, 83]. Over the years, the general nature of research questions in CAM has shifted from efficacy to effectiveness [2, 36, 81]. Pragmatic trials involve randomisation [20, 33, 83, 84] and treatment has to be defined adequately and clearly [53, 83, 85]. In contrast to the wide use of explanatory RCTs addressing efficacy, pragmatic trials have greater external validity [19, 20, 38, 44, 52, 83, 85, 86]. They also allow the evaluation of complex interventions triggering a variety of specific and non-specific effects [29, 36, 87, 88], can include cost evaluations [52, 84], but cannot identify specific mechanisms of action within a treatment [18, 20, 82, 83].

Observational studies might be a feasible method for evaluation of CAM and sometimes lead to results that are comparable with RCTs [15, 17, 20, 44, 52, 59]. This approach could represent a potential alternative if RCTs are seen to be inappropriate, too expensive or too complicated [13, 20, 67, 69, 89], if general effectiveness of an intervention is the focus of interest [24, 52] or to assess CAM use in the population [45]. Results of observational studies can influence the design of further interventional trials [17, 42]. Uncontrolled observational studies, however, give little information about effects of treatment [47], and their weak internal validity must be addressed [20]. A particular method that has been discussed is the Best Case Series [29, 90–92].

The use of quantitative methods, such as factorial and experimental designs [20, 24, 63, 72, 93], has also been proposed. N-of-1 trials (repeated intervention of 1 approach in 1 person) were discussed extensively as a methodology to achieve valid results on the level of the individual patient. It could be appropriate when studying customised treatment of many CAM modalities [20, 49, 72, 74, 94, 95]. However, these trials are uncommon in published research and need to be planned and executed carefully [96].

Qualitative Research Methods
In relation to studies of outcomes of specific therapies, qualitative research may be used to assess the subjective views of individuals [14, 8, 25, 26, 44, 97, 98]. This can help to establish a patient-centred mechanistic understanding of the intervention and its impact, irrespective of whether mechanisms and objective outcomes of treatments are known [16, 26, 56, 97–100]. Qualitative research is unsuitable when trying to establish causal relationships or specific physiological outcomes [101], but is relevant for the investigation of changes in subjective approaches to health and illness [5]. Specific qualitative research methods have been introduced in the literature, such as ethnographic research, interviews and focus groups [5, 16, 98, 102–104]. Case reporting and case studies are particularly valuable to establish complex and contextualised views of the topic under study [67], to gather basic knowledge about CAM treatments [89] or to identify relevant, but uncommon outcomes [89, 105]. However, rigour and sophistication of case reports could be improved [10, 67, 106].

There was a strong consensus that both qualitative and quantitative methods are valuable and should be combined in the CAM research agenda, e.g., qualitative methods to formulate hypothesis on mechanisms (which might be tested by quantitative methods) as well as in specific clinical studies, e.g., to assess reasons for dropouts, identification of the most relevant outcomes or to generally improve interventions [2, 14, 16, 18, 22, 25, 26, 28, 42, 71, 98, 100, 102, 107–110]. The use of qualitative methods has been particularly discussed as a preliminary basis for preparation of clinical trials [25, 28, 29, 79, 97, 101].

Applying Research Methods Used in Conventional Medicine to CAM
Research methods used in conventional medicine can and should be used for research in CAM as well [7, 15, 17, 44, 55, 73, 81, 87, 111–114]. Most authors agreed that the methodological standards of medical research can be applied to CAM research [4, 11, 13, 17, 29, 32, 35, 40, 46, 47, 70, 115, 116], but it might be necessary to adapt the research designs in some areas [6, 12, 15, 43, 57, 58, 62, 69, 87, 88, 117, 118] to account for the complexity of CAM interventions [15, 17, 87, 119, 120]. This is the case not only for CAM, but also for complex and individualised treatments in conventional medicine [72]. However, some authors felt that the underlying assumptions between conventional medicine and CAM differ so fundamentally [8, 18, 39, 64, 121, 122] that specific research methods for CAM are necessary.

Research Priorities
No definite statement can be made concerning the question of which kind of research should be prioritised in CAM, but it was argued that the specification of research priorities is important, as the methods of assessment must be derived from the research question and not vice versa [13]. Various criteria were proposed for deciding on the priorities of future CAM research in general, such as prevalence of use and burden of disease [7, 8, 29, 45, 46, 81, 82, 92, 107, 123–126], and also for specific fields and modalities of CAM [76, 78, 114, 123, 127, 128], where priorities might differ [129]. The context, foundations and philosophical background of CAM treatments [13, 26, 28, 57, 58, 71, 76, 97, 99, 119, 121, 130–132] are an important basis through which to understand the differences between CAM practices and conventional medicine. The safety of different CAM treatments needs to be assessed [46, 57, 58, 64, 82, 89, 91, 119, 133] to protect patients using CAM [46], even though CAM is generally considered safe [55, 81].

There were 2 contradicting views regarding effectiveness versus efficacy studies. Although there seems to be no dis-
The Role of Different Modalities in CAM Research

Issues concerning a broad range of different CAM modalities in CAM research have been discussed in the literature, with acupuncture (as part of Chinese medicine) and homeopathy being the specific CAM modalities addressed most frequently. Use and design of RCTs in acupuncture research have been discussed extensively [3, 5, 32, 40, 43, 48, 50, 55, 151–153]. A major issue is the choice of appropriate control groups (including the design of credible placebo and sham treatments) and blinding [32, 33, 102, 134, 152, 154, 155]. Specific acupuncture-related suggestions for further research have also been given [57, 58, 85, 108, 127, 145]. Similarly, specific issues involving in the design of homeopathic studies have been discussed in detail [75, 112, 113, 156, 157], e.g., the separation of non-specific and specific effects [68, 80] and the handling of patient preferences within a randomisation procedure [88]. The shift from efficacy to effectiveness studies in homeopathy [78, 84] has been suggested to be of more clinical value.

A specific argument that has been raised regarding dietary supplements and herbal medicine is their varying quality and/or composition since there is no adequate standardisation of production for these medicines [30, 55, 158, 159]. Developing an appropriate placebo is crucial especially when there is a difference of taste between the active drug and suggested placebo [49, 99, 158]. There were fewer modality-specific publications for Ayurveda [21], bodywork (such as Feldenkrais) [37], chiropractic [18, 25, 76], classic Arabic medicine [103], diet [73, 120], healing [22, 23, 35, 56, 107], hypnosis [86, 136], traditional Japanese medicine [123], massage [93], meditation [2, 100], Oriental medicine [6], (intercessory) prayer [24, 42, 160], Qigong [51], reflexology [9], Tai Chi [62] and Yoga [115].

Discussion

This literature review summarises and reflects the on-going discussion within the scientific community regarding CAM research over the last 20 years. To the best of our knowledge, this is the first systematic review, following a clearly defined protocol, aimed at assessing the current situation of clinical and epidemiological research methodology in CAM. However, developing definitions of inclusion and exclusion criteria has been proved difficult. Also, although 2 reviewers conducted reference selection and 3 reviewers checked the full texts, first screening was only done by 1 reviewer.

In light of the current literature on CAM research methodology there is broad consensus that the commonly accepted research methods that are used in conventional medicine can and should be applied to evaluate CAM. This applies especially to RCTs. However, the literature reflects a movement from double-blind, placebo-controlled, randomised trials (to explain specific mechanisms and efficacy, as conducted in
drug-research) towards more pragmatic trials that compare meaningful clinical alternatives in heterogeneous groups of patients. Efficacy research was hampered by a lack of consensus-based and testable underlying theories for many CAM modalities, e.g., when designing appropriate placebo or sham treatment. The assumptions underlying the rationale of double-blind placebo-controlled RCTs were also difficult to fulfill for most CAM modalities, e.g., patient and treatment-provider blinding. Consequently, the results of efficacy research have often been inconclusive and difficult to interpret. On the other hand, research into the overall clinical effects of CAM promises more relevant results for clinical decision-making, and within the framework of comparative effectiveness research RCTs of high methodological quality are possible. These challenges and the current trend towards the evaluation of treatments in clinical contexts are not restricted to CAM but affect all areas of complex interventions in medicine [161–163].

Giving priority to comparative effectiveness research does not devalue the importance of basic research on mechanisms of action in CAM, which is needed to facilitate interpretation of efficacy and effectiveness research. A previous independent advisory group [164] stated that trials into effectiveness and cost-effectiveness are primarily needed, but the mechanisms of action of CAM also need to be assessed. In addition, further basic research is needed on the mechanisms of action of placebo intervention or sham controls.

Most authors are in favour of a broad integration of different research methods to gather evidence about the clinical effects of CAM. There is a strong consensus that both qualitative and quantitative methods are valuable and should be combined within the CAM research agenda using a mixed methods approach. This would involve qualitative methodology, for example, to understand the feasibility of running a study, developing the appropriate outcomes and formulating hypotheses about the psychological mechanisms involved in the complex intervention. This information would then be evaluated utilizing quantitative methods in specific clinical studies.

The above-mentioned aspects in clinical and epidemiological CAM research were discussed at a CAMbrella workshop with distinguished experts in the field of CAM research to develop recommendations for further research into CAM. The invited experts were Wayne Jonas, Klaus Linde, Hugh MacPherson, Charlotte Paterson, Harald Walach and Claudia Witt and as members of CAMbrella’s Advisory Board Sean Connolly and Peter Zimmermann. These recommendations form the basis of the CAMbrella ‘roadmap for future clinical and epidemiological research in CAM’.

Disclosure Statement

This project was funded as part of CAMbrella Work Package 7 FP7-HEALTH-2009-3.1–3 (Grant No. 241951).

References