REVIEW

A systematic review of randomised clinical trials of individualised herbal medicine in any indication

R Guo, P H Canter, E Ernst

Postgrad Med J 2007;**83**:633–637. doi: 10.1136/pgmj.2007.060202

See end of article for authors' affiliations

Correspondence to: Peter H Canter, Peninsula Medical School, Universities of Exeter & Plymouth, 25 Victoria Park Road, Exeter, EX2 4NT, UK; peter.canter@pms.ac.uk

Received 27 March 2007
Accepted 21 May 2007

Aim: To summarise and critically evaluate the evidence from randomised clinical trials for the effectiveness of individualised herbal medicine in any indication.
Methods: Search of electronic databases and approaches to experts in the field to identify randomised, controlled clinical trials of individualised herbal medicine in any indication. Independent data extraction and assessment of methodological quality by two authors and best evidence synthesis.
Results: Three randomised clinical trials of individualised herbal medicine were identified. Statistically non-significant trends favouring active over placebo treatment in osteoarthritis of the knee probably result from large baseline differences and regression to the mean. Individualised treatment was superior to placebo in four of five outcome measures in the treatment of irritable bowel syndrome, but was inferior to standardised herbal treatment in all outcomes. Individualised herbal treatment was no better than placebo in the prevention of chemotherapy-induced toxicity.
Conclusions: There is a sparsity of evidence regarding the effectiveness of individualised herbal medicine and no convincing evidence to support the use of individualised herbal medicine in any indication.

Evidence of efficacy for some herbal medicines, but by no means for all those in common use, has increased substantially in the past 20 years.[1] However, most clinical trials of herbal medicine have focused on either standardised extracts of single herbs or standardised formulae reflecting increased sponsorship of such studies by manufacturers in the increasingly important over-the-counter market. The individualised approach, in which patients receive tailored prescriptions comprising a mixture of herbs, is emphasised in most forms of practitioner based herbalism, including European medical herbalism, Chinese herbal medicine and Ayurvedic herbal medicine. The World Health Organization has estimated that 80% of the population in developing countries depends primarily upon herbal medicine for basic health care.[2] Evidence from clinical studies of single herb extracts or standardised formulae cannot be generalised to individualised herbal medicine, and claims by practitioners that the latter has an evidence base requires confirmation. The non-standardised nature of individually prepared herbal prescriptions and the consequent increased potential for adverse events and negative interactions[1] means that safety and effectiveness need to be firmly established before such practices can be endorsed. This systematic review aims to summarise and critically evaluate the evidence from randomised clinical trials for the effectiveness of individualised herbal medicine in any indication. The findings of this review are particularly pertinent because section 12(1) of the UK's Medicines Act relating to regulation of unlicensed herbal remedies made up to meet the needs of individual patients is presently under review.

METHODS

Searching

We searched for randomised clinical trials (RCTs) of any form of individualised herbal medicine in any indication in electronic databases (Medline, Embase, Cochrane Library, CINAHL, AMED) from the inception of the respective database to February 2007 using the search algorithm: [Individual$ OR tailored OR personal$ OR standard$ OR herbal$] AND [Kampo OR herb$ OR plant$ OR Phyto$ OR botanic$ OR extract$ OR (traditional OR Chinese OR herbal OR oriental ADJ medicine)]. No language restrictions were imposed. Additional studies were sought by searching the reference lists of identified trials and reviews, contacting experts in the field who have published similar studies of herbal medicine (n = 5), contacting professional bodies of herbal medicine practitioners, and by hand searching all back issues of the review journal *FACT*. The following 15 professional bodies were contacted: European Herbal Practitioners Association; The Herb Society; The Register of Chinese Herbal Medicine; The College of Practitioners of Phytomedicine; The Herb Society of America; Ayurvedic Practitioners Association; National Institute of Medical Herbalists; The National Herbalists Association of Australia; American Herbalists Guild; American Ayurvedic Association; National Ayurvedic Medical Association; Ontario Herbalists Association; New Zealand Association of Medical Herbalists; Society for Phytotherapy; British Herbal Medicine Association.

Selection of studies

To be included, studies had to be randomised, and controlled. The intervention had to be individualised herbal medicine in which prescriptions were individually tailored for each patient. Studies in any indication were included. Studies combining individualised herbal medicine with other treatments were excluded unless the design allowed the separate evaluation of the effectiveness of the herbal medicine component. The study had to report data on any outcomes for both active and control treatments to be included. An initial assessment against inclusion criteria was made by scanning all titles and abstracts identified by the literature searches (RG). Full text articles of potentially relevant references were retrieved and assessed independently for inclusion by two reviewers (RG, PC). Discrepancies were resolved by discussion between the first two authors (RG, PC) and, if needed, by consulting the third reviewer (EE).

Abbreviations: BSS, Bowel Symptom Score; IBS, irritable bowel syndrome; MYMOP, Measure Yourself Outcome Profile; RCT, randomised clinical trial; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

Quality assessment of studies

Methodological quality of the included trials was assessed using the five point Jadad[3] scale. This is a validated measure of quality of reporting in which points are awarded if the study is described as randomised (+1); the means of carrying out randomisation is described and appropriate (+1); the study is described as double-blind (+1); the means of double-blinding is described and appropriate (+1); and there is a description of withdrawals giving number and reason in both groups (+1). Points are deducted if the method to generate the sequence of randomisation is described and is inappropriate (−1); or if the method of double-blinding is described and is inappropriate (−1).

Data extraction

Data concerning the details of study design, quality of the study, participants, intervention, outcomes and adverse events were extracted independently by two authors (RG, PC) using a pro forma data extraction sheet.

Data analysis

A best evidence synthesis was conducted giving due regard to the quality of studies included. It was anticipated that the data would be clinically heterogenous and unsuitable for meta-analysis and therefore none was planned.

RESULTS

The literature search identified 1345 references. This includes one completed but unpublished trial and two ongoing trials identified through contacts with professional bodies and experts in the field. The initial screening of the titles and abstracts identified 15 potentially relevant references, for which full text articles were obtained for further evaluation. Only three trials were finally included in our review.[4–6] Figure 1 describes the results of the search and inclusion/exclusion process. Agreement between reviewers about study inclusion was 100%.

All three included studies are randomised, double-blind, placebo-controlled RCTs of moderate to good methodological quality. Jadad scores for the three studies were 5,[4] 3[5] and 5.[6] Studies compared: individualised Chinese herbal medicine, standardised Chinese herbal medicine and placebo in irritable bowel syndrome[4]; European individualised herbalism and placebo for osteoarthritis of the knee[5]; and individualised Chinese medical herbalism with placebo for prevention of chemotherapy-induced toxicity in cancer patients.[6] Each RCT is described below and summarised in table 1.

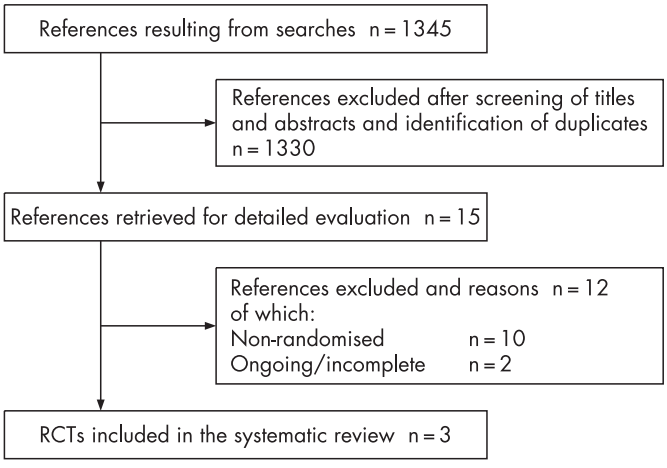

Figure 1 Flow chart of the study selection process. RCTs, randomised controlled trials.

Bensoussan[4] compared individualised Chinese herbal medicine, standardised Chinese herbal medicine and placebo in 116 patients with irritable bowel syndrome (IBS). Treatment lasted 16 weeks and in the individualised group, the prescription could be adjusted by the herbalist at regular intervals. Herbs were administered as encapsulated powders and the standardised treatment was a combination of 20 different herbs. Outcome measures were change in total Bowel Symptom Score (BSS) and global improvement, each assessed separately by the patient and a gastroenterologist, and patient-assessed interference with life. The findings presented in the abstract and results section of this paper differ. The abstract reports statistically significant findings favouring herbal treatment over placebo, but this refers to data derived from standardised and individualised herbal treatment combined together. The results section indicates that there were statistically significant differences favouring standardised treatment over placebo in all five outcome measures, but only four of the five showed significant intergroup differences favouring individualised herbal treatment over placebo. The gastroenterologist's assessments for the main outcome measure, the BSS, were not significantly better than placebo in the individualised group. Overall, changes from baseline and responder rates were larger in the standardised than in the individualised group in all measures. Patient-assessed BSS at a follow up 14 weeks after the end of the trial favoured individualised over standardised treatment, but this difference was not statistically significant.

The data for Hamblin[5] was extracted from a pre-publication draft kindly made available to us by the authors. This study compared 10 weeks of individualised herbal medicine with a placebo tincture in 20 patients diagnosed with osteoarthritis of the knee. The herbal treatments were prescribed by two herbal practitioners each based in a different London general practice surgery. Prescriptions drew upon a formulary of 11 herbs based upon responses to a questionnaire completed by 20 established herbalists. Patients continued with existing pain-killing and anti-inflammatory drugs for the period of the trial and, in addition to the active or placebo treatments, also received dietary advice and daily nutritional supplements consisting of multivitamins and minerals, vitamin C and omega-3 fish oils. Outcome measures were subscale scores and total score for the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) and Measure Yourself Outcome Profile (MYMOP) scores for two symptoms and a daily activity chosen by each patient. Fourteen of the 20 patients enrolled in the study completed the 10 week trial and data analysis is based on these completers. There were no significant differences between groups in changes from baseline for either outcome or their component scores. The authors do report several within-group changes confined to the active treatment group but the only one of these within-group changes to reach statistical significance were the WOMAC stiffness score at 5 weeks and symptom 2 on the MYMOP at both time intervals.

Mok[6] compared the effect of individualised Chinese herbal medicine with that of placebo upon chemotherapy-induced toxicity in patients with early-stage breast and colon cancer. Individualised treatment was prescribed by one of three qualified Chinese herbalists drawing on a stock of 125 different commonly used herbs. Treatments, including the placebo, were dispensed in the form of a herbal tea. Treatment could be adjusted by the herbalist on day 1 and 14 of each cycle of chemotherapy. Chemotherapy was standardised as four 21 day cycles for breast cancer and six 28 day cycles for colon cancer. The trial was terminated early when 50% of the target sample size had been recruited. This was because of a slow accrual rate. Many potential recruits refused the possibility of being randomised to placebo or were already receiving Chinese herbal

Table 1 Included RCTs of individualised herbal medicine

Study	Design (Jadad score)	Participants and indication	Intervention	Outcome measures	Main results	Comments
Bensoussan 1998[4]	Double-blind, placebo controlled RCT (5)	n = 116. 18–75 years (99 analysed). Adults with IBS meeting Rome criteria	16 weeks 1) Individualised CHM 2) Standardised CHM 3) Placebo capsule	1. Total BSS score (4×100 mm VAS for pain/discomfort, bloating, constipation, diarrhoea) a. patient b. gastroenterologist 2. Global improvement a. patient b. gastroenterologist 3. Interference with life assessed by patients (no further details reported)	1a, 1b, 2a, 2b, 3 favoured standardised CHM over placebo 1a, 2a, 2b, 3, 4 favoured individualised CHM over placebo 1b NSD between individualised CHM and placebo 1a, 1b, 2a, 2b, 3 NSD between standardised and individualised CHM	Per protocol analysis for continuous variables Results reported in abstract combine data for the 2 active treatment groups and are inconsistent with those in results section
Hamblin 2007[5]	Double-blind, placebo controlled RCT (3)	n = 20 (14 analysed) Patients with osteoarthritis of the knee	10 weeks 1) Active, individualised herbal medicine prescribed by a herbal practitioner at baseline and week 5 2) Placebo (25% alcohol, caramel colouring, aniseed essential oil flavouring)	At baseline week 5 and 10 1. WOMAC knee health (3×100 mm VAS) a. pain b. physical function c. stiffness d. total 2. MYMOP (3×7-point scales for the preceding 4 weeks) a. symptom 1 b. symptom 2 c. daily activity affected (a, b, c chosen by the patient)	1a, b, c, d. NSD between groups 2a, b, c NSD between groups	Per protocol analysis Success of patient and herbalist blinding not assessed Large baseline differences between groups in WOMAC scores indicating greater severity of osteoarthritis in the active group
Mok 2007[6]	Double-blind, placebo-controlled RCT with randomisation stratified for chemotherapy regimen (5)	n = 120 (111 assessable) Patients with early stage breast or colon cancer receiving chemotherapy (breast cancer—AC 4×3-week; colon cancer—FUFA 6×28 day cycle treated days 1–5)	12 weeks to 6 months 1) Individualised CHM 2) Placebo	1. Haematological toxicity 2. Non-haematological toxicity (16 items) 3. EORTC QoL	1. NSD between groups 2. 1 of 16 items (nausea) favoured active treatment (p = 0.04) 3. NSD between groups	Study terminated early when 50% of target sample recruited, due to difficulty of recruitment

AC, adriamycin and cyclophosphamide; CHM, Chinese herbal medicine; EORTC QoL, European Organisation for Research and Treatment of Cancer Quality of Life questionnaire version 2; FUFA, 5-fluorouracil and folinic acid; IBS, irritable bowel syndrome; MYMOP, Measure Yourself Outcome Profile; NSD, no significant difference; RCT, randomised controlled trial; SS, bowel symptom scale; VAS, visual analogue scale; QoL, quality of life; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

medicine. Data analysis for 111 patients showed no statistically significant differences between groups for the primary outcome measure of haematological toxicity. There were no significant differences between groups for responses to a quality of life questionnaire, and only one of 16 items measuring non-haematological toxicity showed a significant difference favouring the active treatment. This one difference related to nausea, but a similar difference between groups was not observed in the item relating to nausea in the quality of life questionnaire.

DISCUSSION

Summary of main findings

Systematic searches of electronic databases and contacting experts and professional bodies in the field resulted in the location of only three randomised clinical trials of individualised herbal medicine. It should be stressed that professional bodies representing the interests of different practitioner factions from around the world were unable to contribute any more studies than this. In view of the long history and widespread use of medical herbalism, Chinese herbal medicine and Ayurvedic herbal medicine in many and diverse indications, this should be a cause for concern. It indicates that individualised herbal medicine has an extremely sparse evidence base and that there is no convincing evidence supporting its use in any indication. Only one of the three studies[4] indicated that individualised treatment was superior to placebo and this study is particularly important because it found that individualised treatment was inferior to standardised treatment. This study sets a new benchmark for the tailored approach: not only must herbalists demonstrate that individualised treatment is superior to placebo, they must also show, for reasons of cost and safety, that it is superior to standardised treatment. Claims by herbalists who use the individualised approach that their practice is evidence based are disingenuous; this is because evidence supporting the use of herbs for any indication has come almost entirely from the study of single, standardised herbal extracts, not from studies of individualised herbal medicine using combinations of several or many different herbs prepared from inherently variable raw plant materials. The paucity of data supporting the effectiveness of individualised herbal medicine, and the important safety concerns associated with this particular form of phytomedicine, should be taken into account by policymakers concerned with the regulation of practitioners using this modality.

Overall, the results of the three studies included in this review do not provide support for the use of individualised herbal medicine in any indication. Despite optimistic reporting of positive trends in the Hamblin[5] study, there were no statistically significant differences favouring active over placebo treatment in osteoarthritis of the knee, and the trends observed are probably the result of large baseline differences and regression to the mean. While Bensoussan[4] observed that individualised treatment was better than placebo in four of five outcome measures in the treatment of IBS, it was inferior to standardised treatment in all five outcomes and standardised treatment therefore appears to be preferable for reasons of cost and safety. Finally, the Mok study[6] does not provide any convincing evidence that individualised herbal treatment is superior to placebo in the prevention of chemotherapy-induced toxicity. These data indicate that almost all individualised herbal medicine is practised without the support of any rigorous evidence about effectiveness whatsoever.

Limitations of included studies

Although quality of reporting as assessed by the Jadad scale was generally good, all three included studies were characterised by an optimistic interpretation of their findings. In the Bensoussan study,[4] results are presented in rather an obscuring way and may appear to support the use of individualised herbal treatment. The presentation of combined data for the two active treatments in the abstract is particularly misleading. A superficial reading leaves the impression that, overall, the study provides evidence supporting the use of individualised herbal treatment while in fact its most important finding was that individualised treatment was inferior to standardised treatment in all outcomes. Given the additional costs and an increased risk of adverse events resulting from variability of plant material and extracts, species misidentification, contamination and adulteration, and greater potential for negative herb–herb and herb–drug interactions with individualised treatment, the risk-benefit analysis clearly favours standardised treatment. The authors stress the advantage in patient-assessed BSS score at the follow up but this is not statistically significant. Nor is it clear whether the non-significant p value of 0.1 reported in this context refers to a within-group difference or a between group difference, and if the latter, between which two groups.

The authors of the Hamblin[5] study in osteoarthritis of the knee reported trends towards improvement in WOMAC scores and MYMOP scores, which were confined to the active treatment group. This was particularly so for the WOMAC scores where the observed changes of >20% were considered clinically relevant. As this was a feasibility study and was probably underpowered, these data may indicate that a larger RCT would find significant changes favouring active treatment. However, an inspection of the baseline data reveals that there were large differences between the two treatment groups in baseline in all three subscales of the WOMAC and in the total WOMAC score. Patients in the active treatment group appear to have been in more severe pain (44.89 mm vs 27.64 mm on a 100 mm visual analogue scale), to have more stiffness (53.83 mm vs 33.00 mm), and to have more impaired physical function (40.69 mm vs 35.79 mm) and larger total WOMAC scores (42.66 mm vs 33.86 mm). The absence of statistically significant differences at baseline reported by the authors cannot be taken to indicate group comparability. Statistical tests for difference are designed to be conservative and are therefore unsuitable for establishing comparability. The baseline differences seen here are sufficiently large to explain, through regression to the mean, the observed trends appearing to favour active treatment. Hopefully, any future, larger study would have more comparable groups after randomisation. The study design also failed to include any steps to assess the success of blinding. It is important to establish whether or not the placebo tincture is sufficiently convincing to patients to keep them blinded and also to stop them unblinding the medical herbalists during their encounters at week 5 and week 10 of the study.

It is unfortunate that the study of Mok[6] suffered from recruitment problems and was terminated early. This means that it was underpowered in terms of the authors' own calculation. However, there do not appear to be any strong trends in the data favouring active treatment over placebo, other than in the single item of non-haematological toxicity relating to nausea, and this was not confirmed by the corresponding item in the quality of life questionnaire. The nauseas item was one of 16 items relating to non-haematological toxicity and this may be an isolated positive finding occurring as a result of multiple hypothesis testing. The study does not report any testing of the effectiveness of blinding with the placebo tea employed. However, any breakdown in blinding would probably favour active treatment and would not therefore have altered the main findings. The placebo contained *Camellia sinensis* (Chinese Puer tea), *Sojae praeparturum* (black

soybean as seed paste), *Hordeum vulgare* (as sugar from germinated barley seeds) along with food colour and artificial flavour. The authors cite studies supporting a chemopreventative effect of *C sinensis* and there is therefore a theoretical possibility that any effect of active treatment was masked by an anti-toxic effect of the placebo tea.

Implications for future research and clinical practice

All three trials do demonstrate that rigorous RCTs of individualised herbal medicine are entirely feasible. Care should be taken in the choice of a placebo and success of blinding should be measured. Care should be taken to ensure and demonstrate the success of blinding of patients, herbalists and outcome assessors. There is, however, a problem with the generalisability of results from such studies because of the non-standardised nature of the treatment. The large number of single herbs from which individualised treatments are prepared, differences between herbalists in prescribing practice, and the lack of information about the actual treatments prescribed all mean that replication of findings will be made difficult. Even if precise prescribing information was reported for each patient, it is difficult to envisage how these data could be productively used when comparing different studies other than for generating hypotheses about particularly effective component herbs. The lack of standardisation and use of multiple herbs in a single prescription also greatly multiply the safety risks. There are additional risks associated with variability in the diagnostics skills of the practitioner, their awareness or lack of awareness of potential interactions, and their ability or inability to identify red flag symptoms indicating serious diseases requiring immediate mainstream medical treatment. Given the risks and lack of supporting evidence, the use of individualised herbal medicine cannot be recommended in any indication.

Strengths and limitations of this study

Designing a search strategy to locate RCTs of individualised herbal medicine is difficult because of the large number of potential descriptors for such studies, and this is a potential weakness of our systematic review. It is also possible that there are more such studies hidden in the oriental literature which is not adequately indexed by the databases which we have searched. However, these potential weaknesses which may have limited the completeness of our review are, we believe, mitigated by our approaches to experts in the field and to professional bodies of European, American, oriental and Indian herbalists. We know of two other unpublished or incomplete clinical trials of individualised herbal treatment, one for menopausal symptoms and the other for endometriosis. The former, presently under peer review,[7] is a pragmatic pilot trial comparing individualised herbal treatment with waiting list, which will not therefore provide data upon which to base conclusions about the efficacy of treatment. The latter, presently at recruitment stage, will compare active treatment, placebo and waiting list, but again is on the scale of a feasibility study (Andrew Flower, Southampton University, personal communication, 2007).

Comparison with existing literature

This is the first systematic review of RCTs of individualised herbal medicine for any indication.

Conclusion

Individualised herbal medicine, as practised in European medical herbalism, Chinese herbal medicine and Ayurvedic herbal medicine, has a very sparse evidence base and there is no convincing evidence that it is effective in any indication. Because of the high potential for adverse events and negative herb–herb and herb–drug interactions, this lack of evidence for effectiveness means that its use cannot be recommended.

Authors' affiliations

R Guo, P H Canter, E Ernst, Peninsula Medical School, Universities of Exeter & Plymouth, Exeter, UK

Competing interests: None declared

REFERENCES

1 Ernst E, Pittler M, Wider B, eds. *The desktop guide to complementary and alternative medicine, an evidence based approach*, 2nd ed. Mosby Elsevier, 2006.
2 **Vines G**. Herbal harvests with a future: towards sustainable sources for medicinal plants. Plantlife International and WWF-UK report 2004 http://www.plantlife.org.uk/html/goods/documents/HerbalHarvestwithaFuture.pdf.
3 **Jadad AR**, Moore RA, Carrol D, *et al*. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Contr Clin Trials* 1996;**17**:1–12.
4 **Bensoussan A**, Talley NJ, Hing M, *et al*. Treatment of irritable bowel syndrome with Chinese herbal medicine. *JAMA* 1998;**280**:1585–9.
5 **Hamblin L**, Laird A, Walker AF. Reduced osteoarthritis symptoms from practitioner-prescribed herbal treatment: a randomised, double-blind feasibility study (in press).
6 **Mok TS**, Yeo W, Johnson PJ, *et al*. A double-blind placebo-controlled randomized study of Chinese herbal medicine as complementary therapy for reduction of chemotherapy-induced toxicity. *Ann Oncol* 2007;**18**:768–74.
7 **Green J**, Hawkey S, Denham A, *et al*. Changing with herbs: randomised, controlled pilot study of the effectiveness of treatment of menopausal symptoms by medical herbalists. Poster presented to Complementary Medicine and Women's Health: evidence for integration 10 July, 2003.