Effectiveness of herbal medicines for weight loss: A systematic review and meta-analysis of randomized controlled trials

Alison Maunder MPhil1 | Erica Bessell MND1 | Romy Lauche PhD2,3 | Jon Adams PhD2 | Amanda Sainsbury PhD1 | Nicholas R. Fuller PhD1

1University of Sydney, Boden Collaboration for Obesity, Nutrition, Exercise and Eating Disorders, Central Clinical School, Charles Perkins Centre, New South Wales, Australia
2University of Technology Sydney, Australian Research Centre in Complementary and Integrative Medicine (ARCCIM), NSW, Australia
3Department of Internal and Integrative Medicine, Sozialstiftung Bamberg, Germany

Correspondence
Erica Bessell, Boden Collaboration for Obesity, Nutrition, Exercise & Eating Disorders, Charles Perkins Centre D17, University of Sydney, 2006 New South Wales, Australia.
Email: erica.bessell@sydney.edu.au

Funding Information
National Health and Medical Research Council

Peer Review
The peer review history for this article is available at https://publons.com/publon/10.1111/dom.13973.

Abstract

Aim: To update the available evidence on the efficacy and safety of complementary medicines to assist in weight loss by conducting a systematic review and meta-analysis of herbal medicines for weight loss.

Methods: Four electronic databases (Medline, Embase, CINAHL and Web of Science) were searched from inception until August 2018. A total of 54 randomized placebo-controlled trials of healthy overweight or obese adults were identified. Meta-analyses were conducted for herbal medicines with ≥4 studies available. Weight differences of ≥2.5 kg were considered clinically significant.

Results: As a single agent, only *Phaseolus vulgaris* resulted in a statistically significant weight loss compared to placebo, although this was not considered clinically significant. No effect was seen for *Camellia sinensis* or *Garcinia cambogia*. Statistically, but not clinically, significant differences were observed for combination preparations containing *C. sinensis*, *P. vulgaris* or *Ephedra sinica*. Of the herbal medicines trialled in ≤3 randomized controlled trials, statistically and clinically significant weight loss compared to placebo was reported for *Irvingia gabonensis*, *Cissus quadrangularis*, and *Sphaeranthus indicus* combined with *Garcinia mangostana*, among others, but these findings should be interpreted cautiously because of the small number of studies, generally poor methodological quality, and poor reporting of the herbal medicine interventions. Most herbal medicines appeared safe for consumption over the short duration of the studies (commonly ≤12 weeks). Some warrant further investigation to determine effect size, dosage and long-term safety.

Conclusion: There is currently insufficient evidence to recommend any of the herbal medicines for weight loss included in the present review.

KEYWORDS

herbal medicines, meta-analysis, systematic review, weight loss

1 INTRODUCTION

Overweight and obesity has reached epidemic proportions worldwide, with the global prevalence doubling since 1980.1 Obesity affects almost every aspect of health and is strongly associated with type 2 diabetes, cardiovascular disease and various cancers.2-4 Lifestyle interventions, including increasing physical activity and decreasing energy intake, are key components for treating obesity; however, results do not consistently show lifestyle modifications alone as promoting sufficient and sustained weight loss.5,6 Conforming to a new
regimen is problematic for many, and after the initial weight loss, weight regain is common.7

People unable to reduce weight satisfactorily with lifestyle interventions may be candidates for pharmacotherapy as an adjunct. Presently, five medications are approved by the US Food and Drug Administration for chronic weight management,8 only three of which are approved by the European Medicines Association9 or the Australian Therapeutic Goods Administration.10 The addition of these medications to lifestyle intervention increases mean weight loss, however, side effects and costs are often cited reasons for cessation or avoiding use of these drugs.11-13 Many people turn to supplements, which are easier to access and may have fewer side effects, as an alternate approach to maintaining and losing weight. A US study showed that, among people trying to lose weight, 16.1% had used a weight loss supplement in the past year.14

Despite the many supplements available, few have scientific support for their safety and efficacy. Unlike pharmaceutical drugs that require approval before being marketed, clinical evidence is not required for supplements. In some countries, the only requirement is that the supplement contains acceptable levels of non-medicinal substances.15,16 Between 1996 and 2006, 1000 supplements for weight loss were listed on the Australian Register of Therapeutic Goods, but without evaluation of efficacy.17 Sponsors are only required to hold, but not necessarily produce, evidence substantiating their claims. Only 20% of new listings are audited annually to ensure they meet this requirement.17 The increasing popularity of supplements18 further highlights the importance of conducting efficacy and safety studies of weight loss supplements. Many weight loss supplements are based on herbal medicines with a history of international use in health maintenance and disease prevention or treatment, including weight loss.18-20 The World Health Organization (WHO) encourages the integration of herbal medicines of proven quality, safety and efficacy into national healthcare systems.20

There are many small-scale reviews investigating individual herbal medicines for weight loss.21-25 However, it has been 19 years since a comprehensive review of the literature was conducted, combining evidence for all available herbal medicines.26,27 Since then, many new products have been marketed and many more studies have been published, thus the potential benefits and adverse effects of herbal medicines for weight loss require a systematic review to update the body of evidence.17 In direct response, the present paper reports findings from a systematic review and meta-analysis of randomized controlled trials, summarizing and critically evaluating the efficacy and safety of herbal medicines for weight loss.

2 | METHODS

This review was planned, conducted and reported in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines28 and the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions.29

2.1 | Search strategy

Four electronic databases (Medline, Embase, CINAHL and Web of Science) were searched from inception until August 2018. The literature search was constructed around search terms of: 1) obesity or overweight; 2) weight loss; and 3) dietary supplements. These were adapted for each database as necessary. The complete search strategy for Medline is shown as an example in Table S1. Only English-language publications were included.

2.2 | Eligibility criteria

The eligibility criteria for inclusion in this review were as follows:

1) Type of study. Only randomized, parallel-group, placebo-controlled trials (RCTs) were eligible for inclusion. This type of trial reduces bias by controlling for participant differences through randomization and the parallel-group design, and controls for the placebo effect through the use of a control group taking a placebo supplement.

2) Type of participant. Participants aged ≥16 years, of any sex or ethnic origin, with overweight or obesity were included. Overweight and obesity were defined according to WHO classification of body mass index (BMI) ranges. Accordingly, overweight is a BMI range 25 to <30 kg/m² and obesity is BMI ≥30 kg/m², however, in Asian populations the BMI ranges for overweight and obesity are defined as 23 to <25 kg/m² and ≥25 kg/m², respectively.30 Studies were excluded if they included participants with a BMI <25 kg/m² (exclusively or combined with overweight populations, or with a BMI <23 kg/m² for studies conducted in Asian populations), or if the included participants were required to have a specific comorbidity. This did not preclude the possibility that the included participants may have had comorbidities, as is common in populations with overweight or obesity.

3) Types of intervention. RCTs that compared herbal medicines with a placebo were eligible. Herbal medicines included herbs, herbal materials, herbal preparations and finished herbal products, that contain as active ingredients parts of plants, or combinations as defined by the WHO (Table S2). Herbal products were included if they contained, in addition to the main herbal component, natural organic active ingredients,20 such as added caffeine. Studies were excluded if the intervention comprised isolated constituents from plant origin (not the whole plant), plant oils and other dietary supplements (such as fibres or proteins). These interventions will be analysed in a subsequent paper as part of a larger project. Studies were eligible for inclusion regardless of any additional lifestyle interventions used, as long as the same regimen was used in both the herbal intervention and placebo groups. Studies that used a different regimen between the intervention and placebo groups were excluded.

4) Types of outcome measure. Studies with the primary or secondary outcome measuring change in weight or BMI were selected. This was assessed predominantly through the stated objectives
(including the terms weight, BMI, weight loss, anthropometry) or main results presented in each paper.

### 2.3 Study selection

Titles and abstracts were screened by two reviewers (A.M. and R.L. or E.B.) independently. Full texts were imported into Endnote (Version X8, Clarivate Analytics) for papers that were considered to potentially meet the eligibility criteria. These full texts were read by two reviewers (A.M. and R.L. or E.B.) independently and assessed against the above criteria for a final decision about inclusion in the review.

### 2.4 Data extraction

The characteristics of each study were extracted (Table S3), including study details (author, year of publication, country, sample size), participants (age, sex, BMI), intervention (type, dose, duration), and outcome measurements (weight or BMI at baseline and endpoint, and the change in weight or BMI during the study). Safety was assessed between the intervention and placebo groups, including adverse events, treatment-related adverse effects and treatment-related withdrawals.

Quality of reporting was assessed using the Elaborated CONSORT Statement extension for herbal medicine interventions. One reviewer (A.M.) assessed each study for compliance with the checklist, giving a rating of yes or no (Table S4). The type of herbal medicine formulation and placebo used were also recorded for each study.

### 2.5 Assessment of risk of bias

Two reviewers (A.M. and R.L. or E.B.) independently assessed risk of bias using the standard risk-of-bias assessment tool for parallel-group trials, as recommended by the Cochrane Collaboration. This tool assesses risk of bias in random sequence generation and allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessors (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and other biases. Other biases include other major flaws in the trial design or methodology or reporting, conflicts of interest, or differing levels of adherence to the added lifestyle interventions between the intervention and placebo groups (if reported). Each study was assessed on these seven criteria, giving a rating of “low risk”, “high risk”, or “unclear”. Studies were considered to have a low risk of bias for a criterion if the report detailed an adequate methodology that did not introduce potential bias, as per Table 8.5.d of the Cochrane Handbook. Likewise, studies were considered to have a high risk of bias for a criterion if the report detailed a methodology that is known to be inadequate in removing potential bias. Studies were rated as having an unclear risk of bias if there was no information or insufficient information provided to judge the criterion otherwise.

### 2.6 Data analysis

#### 2.6.1 Assessment of overall effect size

Separate meta-analyses were conducted for each herbal medicine for comparisons of weight only (using absolute weight at endpoint or weight loss) between herbal medicine interventions and placebo. If these weight values were not reported, attempts were made to obtain the missing data from the authors of the RCT directly. If there was no further information provided, these studies were excluded from the meta-analysis and risk of “other bias” was considered high. If at least four studies of a herbal medicine assessing this specific outcome were available, meta-analyses were conducted using Review Manager 5 software (Version 5.3.5, Copenhagen, The Cochrane Collaboration) by random-effects models.

Mean differences (MDs) or standardized mean differences (SMDs) with 95% confidence intervals (CIs) were calculated between the herbal medicine and placebo. SMD was reported in the meta-analyses when the available study results were in varying units of weight. Where no standard deviations were available, they were calculated from standard errors, CIs or t-values. A negative MD or SMD indicated beneficial effects (ie, greater weight loss) of the herbal medicine compared to the placebo. Clinical significance was defined as a 2.5-kg or more difference between herbal medicine and placebo controls at endpoint. Weight loss of this amount was found to be sufficient to reduce the incidence of cardiovascular events and mortality in a recent large-scale clinical trial. For each herbal medicine with adequate studies identified, two meta-analyses were conducted to separate the studies testing single-herb products and studies testing combination preparations containing other active herbs in addition to the herb of interest.

#### 2.6.2 Assessment of heterogeneity

Heterogeneity statistics are shown in the legend of each figure containing a forest plot. \( \tau^2 \) estimates the variance of the true effect size, while the \( P \) value represents the significance of this estimate. \( I^2 \) indicates what proportion of the observed variance in effect sizes between studies is attributable to variability in the true effect size. The remainder of the variability is thus attributed to sampling error. For instance, an \( I^2 \) of 0% indicates that all the observed differences between study effect sizes are attributable to sampling error; an \( I^2 \) of 100% indicates that all the observed differences between studies are attributable to variability in the true effect. \( I^2 \) values of 25%, 50% and 75% are generally considered low, moderate and high, respectively.
3 | RESULTS

3.1 | Identification of studies

A total of 597 references were identified (Figure 1), including 228 identified via Medline, 193 via Embase, 138 via Web of Science, and 38 via CINAHL. After screening, 54 clinical trials met the criteria for inclusion in this review.

3.2 | Risk of bias within studies

The judgements about each risk-of-bias domain are presented as percentages across all trials in Figure 2, and the judgements for each trial are detailed in Figure S1. Most trials provided limited information about study design and methodology. Risk of selection bias was mainly unclear: only 20 of 54 RCTs (37%) adequately reported randomization. Only six RCTs (11%) reported adequate allocation concealment. Performance bias was mainly unclear (98%), with only one study reporting the procedure of blinding of participants or study personnel. Detection bias was also mostly unclear (91%), with two studies being considered low risk, while the other three studies were considered high risk. Attrition bias was low in 26 studies (48%), but was judged as high in 16 studies (30%). The risk of selective outcome reporting was unclear in most studies (78%), and only four studies (7%) had prospectively registered their trial in a public trial registry. Forty-two studies (78%) were considered to have a high risk of "other bias". Common reasons included failure to mention whether or not
the authors had conflicts of interest, and insufficient detail of the herbal medicine intervention according to the elaborated CONSORT checklist items for reporting of RCTs of herbal medicine (Table S4).

3.3 | Study characteristics

The characteristics of the studies are summarized in Table S3. The herbs investigated most frequently were *Camellia sinensis* (n = 12, 22%), *Garcinia cambogia* (n = 11, 20%), *Phaseolus vulgaris* (n = 7, 13%) and *Ephedra sinica* (n = 5, 9%).

The same lifestyle regimen was implemented in both the intervention and placebo groups in each study. Participants were asked to maintain their usual diet in 24 studies (44%), were provided with dietary advice or meal plans for general health and/or weight loss in 23 studies (43%), or were provided with a weight maintenance diet through food deliveries (n = 4) or inpatient settings (n = 2). The remaining study provided participants with low-energy meal replacement shakes. In terms of physical activity, 38 studies (70%) requested participants to maintain their usual levels, 11 studies (20%) provided exercise advice, and five studies (9%) held exercise sessions for participants to attend.

3.4 | *Camellia sinensis*

*Camellia sinensis* or green tea was investigated in 12 studies (n = 1179), seven as single-herb products,34-40 four as combination preparations41-44 and one as part of a traditional Chinese medicine.45 Dosages of *C. sinensis* were reported in six studies, ranging from 200 mg to 2400 mg daily. More studies reported on the active ingredients, including epigallocatechin gallate (n = 7; 90 mg to 857 mg...
daily) and caffeine (n = 7; 27 mg to 225 mg daily). There were no similarities between the formulas of the combination preparations. The study durations ranged from 8 to 13 weeks.

There was inadequate reporting of trial design and methodology across the studies, which largely prevented assessment of risk of selection bias, performance bias and detection bias. Six studies were at low risk of attrition bias, while four studies were rated as high risk. Seven studies were rated as high risk from other biases for reasons including lack of trial registration, potential conflict of interest as some authors worked for the company owning the herbal medicine, and trials including a run-in period to remove participants with low compliance.

Meta-analysis of seven studies (n = 654) investigating *C. sinensis* as a single herb showed no benefit over placebo (−0.27 kg; 95% CI −0.73, +0.18; P = 0.24) [Figure 3]. Meta-analysis of five studies (n = 393) investigating combination preparations containing *C. sinensis* showed a statistically significant effect of −1.63 kg (95% CI −2.41, −0.85; P < 0.001) compared to placebo. This weight difference is not considered clinically significant as it is below the benchmark of −2.5 kg.11

In these two meta-analyses, the single-herb studies were similar to each other (Tau² = 0.00, P = 0.83; I² = 0%), while the studies of combination preparations were more heterogeneous (Tau² = 0.25, P = 0.19; I² = 35%) potentially because of the variability in the herbal medicine formula and the greater representation of male participants.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean (SD) Total</th>
<th>Mean (SD) Total</th>
<th>Mean Difference (IV, Random, 95% CI)</th>
<th>Mean Difference (IV, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heymsfield 1998</td>
<td>−3.2 (3.3) 42</td>
<td>3.9 (2.4%) 42</td>
<td>24.8% 0.25 (−0.18, 0.68)</td>
<td>0.25 (−0.18, 0.68)</td>
</tr>
<tr>
<td>Kim 2011</td>
<td>−0.68 (1.83) 29</td>
<td>−0.65 (2.32) 29</td>
<td>21.7% −0.01 (−0.53, 0.50)</td>
<td>−0.01 (−0.53, 0.50)</td>
</tr>
<tr>
<td>Mattes 2000</td>
<td>−3.7 (3.1) 42</td>
<td>−2.4 (2.9) 47</td>
<td>25.1% −0.43 (−0.85, −0.01)</td>
<td>−0.43 (−0.85, −0.01)</td>
</tr>
<tr>
<td>Preuss 2004a</td>
<td>83 (21.5035) 10</td>
<td>86 (15.633) 10</td>
<td>12.1% −0.15 (−1.03, 0.73)</td>
<td>−0.15 (−1.03, 0.73)</td>
</tr>
<tr>
<td>Preuss 2004b</td>
<td>87.21 (14.0792) 19</td>
<td>78.84 (9.2) 16</td>
<td>16.4% 0.68 (−0.01, 1.36)</td>
<td>0.68 (−0.01, 1.36)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>142</td>
<td>144 (100.0%) 100</td>
<td>Heterogeneity: Tau² = 0.10; Ch² = 9.12, df = 4 (P = 0.06); I² = 56%</td>
<td>0.04 (−0.33, 0.41)</td>
</tr>
</tbody>
</table>

Test for overall effect: Z = 0.22 (P = 0.82)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Mean (SD) Total</th>
<th>Mean (SD) Total</th>
<th>Mean Difference (IV, Random, 95% CI)</th>
<th>Mean Difference (IV, Random, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chong 2014</td>
<td>−2.26 (2.37) 46</td>
<td>−0.56 (2.34) 45</td>
<td>15.5% −0.72 (−1.14, −0.29)</td>
<td>−0.72 (−1.14, −0.29)</td>
</tr>
<tr>
<td>Opala 2006</td>
<td>−2 (2.6) 47</td>
<td>−1.5 (3.5) 51</td>
<td>15.7% −0.16 (−0.56, 0.24)</td>
<td>−0.16 (−0.56, 0.24)</td>
</tr>
<tr>
<td>Preuss 2004a</td>
<td>80.8 (15.59) 10</td>
<td>86 (15.633) 10</td>
<td>12.2% −0.32 (−1.20, 0.56)</td>
<td>−0.32 (−1.20, 0.56)</td>
</tr>
<tr>
<td>Preuss 2004b</td>
<td>86.72 (15.5705) 18</td>
<td>78.84 (9.2) 16</td>
<td>13.7% 0.59 (−0.10, 1.28)</td>
<td>0.59 (−0.10, 1.28)</td>
</tr>
<tr>
<td>Thom 2000</td>
<td>−3.5 (2) 20</td>
<td>−1.3 (1.4) 20</td>
<td>13.7% −1.25 (−1.93, −0.57)</td>
<td>−1.25 (−1.93, −0.57)</td>
</tr>
<tr>
<td>Toromanyan 2007</td>
<td>−4.167 (2.7167) 30</td>
<td>−0.55 (1.0001) 28</td>
<td>14.3% −1.72 (−2.33, −1.11)</td>
<td>−1.72 (−2.33, −1.11)</td>
</tr>
<tr>
<td>Vasques 2008</td>
<td>0.38 (1.9) 32</td>
<td>−0.48 (1.7) 26</td>
<td>14.9% 0.47 (−0.06, 0.99)</td>
<td>0.47 (−0.06, 0.99)</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>203</td>
<td>196 (100.0%) 100</td>
<td>Heterogeneity: Tau² = 0.54; Ch² = 46.00, df = 6 (P &lt; 0.00001); I² = 87%</td>
<td>−0.44 (−1.03, 0.15)</td>
</tr>
</tbody>
</table>

Test for subgroup differences: Ch² = 1.86, df = 1 (P = 0.17), I² = 46.3%
Safety was reported in nine of the 12 studies, including eight reporting on adverse events and six reporting on treatment-related adverse effects. Of the eight studies reporting on adverse events, four reported none,\(^{35,36,40,44}\) three reported few adverse events, with adverse effects. Of the eight studies reporting on adverse events, four reporting on adverse events and six reporting on treatment-related effects. Malabar tamarind was examined in 11 studies.\(^{33,34}\) Garcinia cambogia was investigated in 11 studies.\(^{35,37,43,49,50}\) Ten of the 11 studies investigating G. cambogia safety, including five reporting on adverse events and seven reporting on treatment-related adverse effects. Of the five studies reporting on adverse events, all reported no significant differences in adverse events between the intervention and placebo groups.\(^{38}\) Of the six studies reporting on treatment-related adverse effects, three reported more effects in the intervention compared to placebo groups.\(^{35,37,43}\) One reported more effects in the placebo compared to intervention group,\(^{38}\) and two reported no treatment-related adverse effects.\(^{41,44}\)

### 3.5 | Garcinia cambogia

Garcinia cambogia or Malabar tamarind was examined in 11 studies (n = 967), six as single-herb products\(^{46-51}\) and five as combination preparations.\(^{41,43,52-54}\) Dosages of G. cambogia were reported in nine studies, and ranged from 300 mg to 4667 mg daily, delivered over 8 to 17 weeks. Most studies also reported the daily dose of the active ingredient, hydroxycitric acid, ranging from 1200 mg to 2800 mg. There was no similarity between the formulas of the combination preparations.

Inadequate reporting in the studies increased risk of selection bias, performance bias and detection bias. Risk of attrition bias was low in six studies (0% to 18% attrition) and high in three studies (29% to 47% attrition). Five studies were at high risk of reporting bias, often because no absolute values were reported, only graphical representations or differences from baseline to endpoint. Ten of the 11 studies were rated as at high risk of other biases because of absence of trial registration and lack of detail regarding the herbal medicine.

Meta-analysis of five studies (n = 285) of G. cambogia as a single herb produced a non-significant effect on weight, with an SMD +0.04 (95% CI −0.33, +0.41; P = 0.82) compared to placebo (Figure 4). One study was excluded from the meta-analysis because of incomplete data. This study reported a significantly greater reduction in BMI in those taking G. cambogia compared to placebo.\(^{33}\) Combination preparations containing G. cambogia also had a non-significant effect, with SMD −0.44 (95% CI −1.03, +0.15; P = 0.14) compared to placebo in 399 participants. Overall, the single-herb studies were considered to have a moderate level of heterogeneity (\(\tau^2 = 0.10, P = 0.06; I^2 = 56\%\)), while studies of combination preparations were observed to have significant heterogeneity (\(\tau^2 = 0.54, P < 0.001; I^2 = 87\%\)).

Ten of the 11 studies investigating G. cambogia reported safety, including five reporting on adverse events and seven reporting on treatment-related adverse effects. Of the five studies reporting on adverse events, all reported no significant differences in adverse events between the intervention and placebo groups.\(^{41,43,46,49,50}\) Of the seven studies reporting on treatment-related adverse effects, five reported no effects,\(^{41,47,51-53}\) one reported no difference between the intervention and placebo groups,\(^{54}\) and one reported increased unpleasant gastrointestinal effects of the intervention compared to placebo (affecting 8 vs 0 participants).\(^{43}\)

### 3.6 | Phaseolus vulgaris

Phaseolus vulgaris or white kidney bean was investigated in seven studies (n = 531), four as single-herb products\(^{35-38}\) and three as combination preparations.\(^{43,52,59}\) Six studies reported dosages ranging from 445 mg to 3000 mg daily over 4 to 13 weeks. There was little similarity in the formulas of the combination preparations.

A lack of detail in trial design and methodology was common across the studies. Risk of attrition bias was low in five studies (0% to 18% attrition). Five studies were rated as being at high risk of other biases because of absence of trial registration and lack of detail of the herbal medicine; two studies also included a run-in period to remove participants with poor compliance.

A statistically but not clinically significant effect on weight was found for P. vulgaris as a single herb compared to placebo in 277 participants (−1.61 kg; 95% CI −1.96, −1.26; P < 0.001 [Figure 5]). One study was excluded from the meta-analysis due to incomplete data. This study did not report final weights nor provide standard deviations for weight losses, however, it did note no statistically significant difference in the intervention compared to placebo.\(^{37}\) Combination preparations containing P. vulgaris had a statistically significant effect on weight (−1.85 kg; 95% CI −3.24, −0.46; P = 0.009) compared to placebo in 190 participants (Figure 5); again, this did not meet the
benchmark for clinical significance of −2.5 kg.11 The single-herb studies were found to be very similar to each other (Arr2 = 0.00, P = 0.42; I² = 0%) while the studies of combination preparations were significantly heterogeneous (Arr2 = 1.08, P = 0.03; I² = 72%).

All seven studies that investigated P. vulgaris reported on treatment-related adverse effects, and three additionally reported on adverse events, which were minimal and comparable between the intervention and placebo groups.43,56,57 Five studies reported no treatment-related adverse effects,52,55–58 and the remaining two studies reported increased gastrointestinal effects, such as discomfort and flatulence, in the intervention compared to the placebo group.43,59 The flatulence was reported to resolve within 1 to 3 days.59 In one study, a participant from the intervention group withdrew due to constipation.59

### 3.7 | Ephedra sinica

**Ephedra sinica** or ephedra was examined in five studies (n = 546), one as a single-herb product60 and four in combination preparations,61–63 including one Japanese medicine.64 Daily dosages of E. sinica were reported in only two studies (125 mg and 200 mg), while all reported a daily dose of the active ingredient, ephedrine, ranging from 24 mg to 90 mg. Apart from caffeine, which was added to all the combination preparations (25 mg to 280 mg daily), there was little similarity between their formulas. The study durations ranged from 8 to 26 weeks.

Four of the five studies reported an appropriate method for random sequence generation, but no studies adequately reported allocation concealment or blinding methods, so no assessments could be made for these criteria. Three studies had a high risk of attrition bias (28% to 54% attrition) and the remaining two had a low risk (5% and 16% attrition). All five studies were found to be at high risk from other biases as none were registered and some did not adequately report on adverse events or did not provide P values for baseline or outcome comparisons.

Only one study examined E. sinica alone, so no meta-analysis was performed. Meta-analysis of four studies (n = 382) of combination preparations containing E. sinica showed a statistically but not clinically significant effect with SMD −0.58 (95% CI −0.78, −0.37; P < 0.001; Arr2 = 0.00, P = 0.52; I² = 0%) compared to placebo (Figure 6).

All five studies that investigated E. sinica reported on safety outcomes, including one reporting on adverse events, four reporting on treatment-related adverse effects, and one reporting on withdrawals only. In the one study reporting on adverse events, a total of 196 adverse events were recorded, with no difference between intervention and placebo (77% vs 76% of participants affected).63 Of the four studies reporting on treatment-related adverse effects, all observed more effects in the intervention group than the placebo: 44 vs 8, 20 vs 15, 60 vs 43, and 30 vs 22.60–63 Symptoms included palpitations, hypertension and various gastrointestinal side effects. Three studies noted withdrawals because of treatment-related adverse effects, which were higher in the ephedra group compared to placebo overall: 3 vs 0,64 7 vs 8,61 and 8 vs 0.62

### 3.8 | Irvingia gabonensis

**Irvingia gabonensis** or African mango was investigated in three studies (n = 232), two as single-herb products65,66 and one combined with Cissus quadrangularis.67 All studies were performed by the same research team in Cameroon over 4 to 10 weeks. Daily dosage of I. gabonensis was reported in two studies as 300 mg and 3150 mg. Studies inadequately reported trial design and methodology, and in some cases, outcomes, which prevented assessment of risk of bias across most criteria. For this reason, all studies were rated as high risk from other biases. Notably, one study reported a difference in mean body weight at baseline of >25 kg between the intervention and placebo group,66 and another reported a difference of >1600 kJ in mean energy intake during the study between the intervention and placebo group.65 All three studies reported a statistically significant effect on weight between −4.0 kg and −12.1 kg compared to placebo, which is above the benchmark for clinical significance of −2.5 kg. Two of the three studies reported on safety, recording no significant differences in adverse events between the intervention and placebo groups.65,67

### 3.9 | Ilex paraguariensis

**Ilex paraguariensis** or yerba mate was investigated in three studies (n = 182) over 6 to 12 weeks, one as a single-herb product68 and two in combination preparations.43,69 Daily dosage of I. paraguariensis was reported in two studies as 336 mg and 3150 mg. All studies were rated as having low risk of attrition bias but a high risk of other biases because of a lack of detail of the herbal medicine or discrepancies in reporting of results. Two studies were rated as being at high risk of reporting bias as the results were only displayed in graphs and lacked absolute values or P values. Of the three studies, only one study on a combination preparation69 reported a statistically and clinically significant weight difference of −4.8 kg compared to placebo. Safety was reported in two trials,43,68 with one withdrawal as a result of a serious adverse event in the placebo group.68 Both trials reported no differences between the intervention and placebo groups for adverse events. One trial additionally reported that blood cell count, markers of liver and kidney function, and vital signs remained within the normal range in both the intervention and placebo groups.68

### 3.10 | Cissus quadrangularis

**Cissus quadrangularis** or Veld grape was examined in two studies (n = 164). One study had two arms, investigating C. quadrangularis either as a single herb or combined with I. gabonensis,67 and the
other study examined it combined with C. sinensis.\textsuperscript{42} Both studies were performed by the same research team in Cameroon over 8 to 10 weeks, and were rated as having an unclear risk of bias for most domains because of inadequate reporting of the trial design and methodology. Both studies were rated at high risk from other biases for this reason, and one study inadequately reported baseline characteristics.\textsuperscript{67} All studies reported a statistically and clinically significant (at least −2.5 kg) effect on weight between −4.3 kg and −9.8 kg compared to placebo. Safety was recorded in both trials and the number of adverse events did not differ between the intervention and placebo groups.

### 3.11 Glycyrrhiza glabra

Glycyrrhiza glabra or licorice root was investigated in two studies (n = 149), one as a single-herb product over 8 weeks\textsuperscript{70} and one in a Japanese medicine over 24 weeks.\textsuperscript{64} Both studies reported an appropriate method for random sequence generation, thus were rated low risk in this criterion. Both studies were rated at high risk from other biases because of a lack of detail about the herbal medicine. A statistically and clinically significant weight difference of −3.9 kg compared to placebo was reported only for the combination preparation.\textsuperscript{64} As this product contained multiple ingredients, it is not possible to attribute the effects to G. glabra alone. Safety was recorded in both studies, with the single-herb study reporting no treatment-related adverse events.\textsuperscript{70} The combination-herb study reported three withdrawals in the active group, as participants experienced loose bowel movements, versus no withdrawals in the placebo.\textsuperscript{64} These adverse events could be due to G. glabra or the other active ingredients.

### 3.12 Sphaeranthus indicus and Garcinia mangostana

Sphaeranthus indicus (East India globe thistle) combined with Garcinia mangostana (mangosteen) was examined in two studies (n = 120) conducted by two different research groups in India.\textsuperscript{71,72} Daily dosages were 600 mg of S. indicus and 200 mg of G. mangostana taken over 8 and 16 weeks. One study was rated at low risk of selection bias, attrition bias and other biases; however, it had a high risk of detection bias as unblinding occurred prior to data analysis.\textsuperscript{72} The other study was rated as having a low risk of attrition bias and reporting bias but could not be assessed for the other criteria because of inadequate reporting.\textsuperscript{71} Both studies reported statistically and clinically significant weight differences of −3.7 kg\textsuperscript{71} and −4.0 kg\textsuperscript{72} compared to placebo. Safety data were collected in both studies, which reported no significant differences in adverse events between the active and placebo groups. One study reported no significant differences in biochemical markers,\textsuperscript{71} and the other study reported no treatment-related adverse events.\textsuperscript{72}

### 3.13 Evidence from single RCTs for herbal medicine preparations

Eight single-herb products and nine combination preparations were investigated in only one RCT each. The duration of the 17 studies ranged from 6 to 16 weeks, with a total of 1168 participants. Apart from five studies which reported an appropriate method for random sequence generation, there was a lack of detail in the reporting of trial design and methodology, largely preventing the assessment of risk of selection bias, performance bias and detection bias. Risk of attrition bias was low in seven studies but high in six studies. Risk of reporting bias was high in three studies and 14 studies had a high risk of other biases, including eight with no trial registration, five with insufficient detail about the herbal medicine, and five with potential conflicts of interest as authors held positions or shares in the herbal medicine company or the company had input in trial design.

Statistically significant weight differences favouring the intervention compared to placebo were reported in the RCTs investigating Aster spathulifolius Maxim,\textsuperscript{73} Evodia rutaecarpa,\textsuperscript{60} Garcinia atroviridis,\textsuperscript{74} Gynostemma pentaphyllum,\textsuperscript{75} Zingiber officinalne Roscoe,\textsuperscript{76} Punica granatum,\textsuperscript{77} Dolichos biflorus and Piper betle,\textsuperscript{78} Scutellaria baicalensis and Platycodon grandiflorum,\textsuperscript{79} Imperata cylindrica,\textsuperscript{80} Lippia citriodora\textsuperscript{81,82} and the traditional herbal medicine Triphala.\textsuperscript{83} Four of these studies reported a weight difference compared to placebo which met the benchmark for clinical significance of −2.5 kg, ranging between −2.5 kg and −3.5 kg.\textsuperscript{73,77,78,83} No effect on weight was found for Brassica rapa L,\textsuperscript{84} Caralluma fimbriata,\textsuperscript{85} Glycine max,\textsuperscript{47} Citrus aurantium and Hypericum perforatum,\textsuperscript{86} Magnolia officinalis and Phellodendron amurense\textsuperscript{87} and the traditional herbal medicine Taeeumjowi-tang.\textsuperscript{88}

Safety data were collected in all trials, including reports on adverse events (100%), liver function tests,\textsuperscript{47,60,73,80,83,84,88} renal function tests,\textsuperscript{60,73,80,88} and serum biochemistry,\textsuperscript{60,73,75,86} with no differences from placebo reported. Ten trials reported on treatment-related adverse events, and none were evident.

### 4 DISCUSSION

This review identified 54 RCTs of herbal medicines for weight loss, including a total of 4331 participants. While the meta-analyses demonstrated that some herbal medicines have statistically significant effects on weight compared to placebo, the weight loss was not clinically relevant. These herbs were P. vulgaris as a single agent, and combination preparations containing C. sinensis or P. vulgaris. No statistically significant effects were seen for G. cambogia. These herbs thus cannot currently be recommended for weight loss, which is similar to the conclusions of a 2012 Cochrane review of C. sinensis,\textsuperscript{89} a 2011 systematic review of G. cambogia\textsuperscript{90} and a 2011 systematic review of P. vulgaris.\textsuperscript{91} Statistically and clinically significantly greater weight losses were seen for seven other herbal medicines compared to placebo, but each of these were investigated in three or fewer RCTs and were therefore too few to be meta-analysed. The results
need to be interpreted cautiously as nearly all RCTs included in this review were found to be at high risk of bias in at least one domain, notably because of insufficient detail on the herbal medicine and inadequate reporting of the trial design and methodology. While this high risk of bias would decrease confidence in our conclusions if they favoured herbal medicines, most herbs were not found to have an unequivocal beneficial effect on weight, and so the risk of attributing a positive effect when none exists is not present.

Safety data were scarce in the reports of RCTs underpinning this review. Of the limited data available, our findings suggest that G. cambogia is safe, consistent with a previous systematic review.90 More recently, however, formulations containing the active ingredient of G. cambogia, hydroxycitric acid, have exhibited hepatotoxicity.92 A review concluded that up to 2800 mg of hydroxycitric acid daily was safe,93 but this was based on small clinical trials conducted over 12 weeks. Further studies are required to determine the long-term safety of G. cambogia. Additionally, the safety of E. sinica, which was found to produce a statistically and clinically significant weight reduction compared to placebo in two of five RCTs, remains controversial. While the adverse effects of E. sinica are well established in the literature, it is an important traditional Chinese medicine used for asthma and respiratory decongestion.94 In traditional use, it is administered as a decoction, rarely combined with caffeine, and the dosage is typically limited to 7 to 10 days.95 E. sinica contains ephedrine alkaloids which act as sympathomimetics and can lead to pronounced cardiovascular and central nervous system-stimulating effects. The addition of caffeine enhances these effects.95 In 2004, the US Food and Drug Administration banned supplements containing ephedrine alkaloids of the ephedrine type in response to scientific evidence and adverse event reports, including deaths.96 Likewise, its use has been restricted in Germany, Israel, Canada, Australia and New Zealand.97,98 Given this, E. sinica is not recommended as a weight loss supplement. It should be noted that rigorous collection and reporting of safety as required in phase III trials for pharmaceuticals is not required prior to marketing herbal medicines, and so the safety of herbal medicines in general, with the exception perhaps of E. sinica, cannot be ascertained from the present review.

Three of the seven herbal medicines alluded to above that demonstrated statistically and clinically significantly greater weight losses than placebo were C. quadrangularis, S. indicus combined with G. mangostana, and I. gabonensis. However, these effects were observed in RCTs of only 8 to 10 weeks’ duration, and all of the two to three studies conducted on each of these herbal medicines were by the same research groups, which introduces a high risk of selective reporting and confirmation bias, therefore, more evidence is needed before a definitive decision on their effect on weight loss can be reached. This conclusion is consistent with a 2017 systematic review of C. quadrangularis,99 and a 2013 systematic review of I. gabonensis.100 In contrast, our review revealed inconsistent results among the few studies on I. paraguariensis and G. glabra, perhaps due to the varied combination preparations studied. A. spathulifolius as a single-herb product and combination preparations with P. granatum and Undaria pinnatifida, P. betle and D. biflorus, as well as the traditional herbal medicine Triphala, induced clinically significant weight losses compared to placebo over 8 to 16 weeks. Despite the positive results in most studies, a greater number of larger more rigorous studies are needed to evaluate the efficacy of these herbs.

This systematic review is the first in 19 years to comprehensively review the literature and provide meta-analyses of the available evidence from RCTs of herbal medicines for weight loss. Across the 54 studies, a wide range of countries and cultures were represented, improving the generalizability of the findings. However, despite the introduction in 2006 of the Elaborated CONSORT Statement extension for herbal medicine interventions,31 many studies were still not adequately reporting on trial design and methodology, or on the composition of the herbal medicines being investigated.

The findings of this review are only applicable in the contexts in which the studies were conducted and are limited by their quality and variability. Additionally, the review included all RCTs of herbal medicines for weight loss, regardless of the additional lifestyle interventions delivered, providing that the same lifestyle intervention was used in both the intervention and placebo groups. These factors may have contributed to the heterogeneity observed in the effect sizes. Another limitation of the current review is that only English-language publications were included, which may restrict the representation of research from non-English speaking countries and limit the generalizability of the findings. Any non-published negative studies were naturally not identified; however, these studies would not alter the conclusions as no herbal medicines were found to be unequivocally beneficial for weight loss.

In conclusion, many of the herbal medicines meta-analysed produced statistically significant weight loss compared to placebo, but these effects were below the level of weight loss that is considered clinically important. Some herbal medicines warrant further investigation in larger more rigorous studies to determine the effect size, dosage and long-term safety, notably C. quadrangularis, S. indicus and I. gabonensis, and some products were only represented by one study. Many of the included studies were small, of poor design and methodological quality, with inadequate reporting of the herbal medicine interventions. Future RCTs would benefit from trial registration, and ensuring the study is conducted and reported in a way that minimizes bias and conforms with the CONSORT Statement for reporting of clinical trials.31 Currently there is insufficient evidence to recommend any of these herbal medicines for weight loss.

ACKNOWLEDGMENTS

This work was supported by the National Health and Medical Research Council (NHMRC) of Australia via a Senior Research Fellowship to A.S. (1135897).

CONFLICTS OF INTEREST

A.S. is the author of The Don’t Go Hungry Diet (Bantam, Australia and New Zealand, 2007) and Don’t Go Hungry For Life (Bantam, Australia and New Zealand, 2011); and has provided paid presentations at conferences for Eli Lilly, the Pharmacy Guild of Australia, Novo Nordisk, the Dietitians Association of Australia, Shoalhaven
Family Medical Centres, the Pharmaceutical Society of Australia, and Metagenics. A.S. served on the Nestlé Health Science Optifast® VLCD™ Advisory Board from 2016 to 2018. N.R.F. is the author of *Interval Weight Loss* (Penguin Random House; 2017) and *Interval Weight Loss For Life* (Penguin Random House; 2018), and has received research grants for clinical trials funded by SFI Research, the Australian Eggs Corporation, Sanofi-Aventis, Novo Nordisk, Allergan, Roche Products, MSD, and GlaxoSmithKline. A.M., E.B., R.L. and J.A. have no conflicts to declare.

A.M., R.L. and N.F. contributed to the conception of the review and the development of the protocol. A.M., R.L. and E.B were involved in data extraction. A.M. analysed the data. A.S. mentored A.M. and E.B. in interpretation of the results and preparation of sequential drafts of the manuscript. All authors reviewed and contributed to the manuscript drafted by A.M. and E.B. All authors read and approved the final version.

**ORCID**

Erica Bessell [https://orcid.org/0000-0002-5938-9917](https://orcid.org/0000-0002-5938-9917)

Nicholas R. Fuller [https://orcid.org/0000-0002-6823-5872](https://orcid.org/0000-0002-6823-5872)

**REFERENCES**


47. Mitraheri E, Namazi N, Alizadeh M, Sargheini N, Karimi S. Effects of dried licorice extract with low-calorie diet on lipid profile and


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.