The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/79686

Please be advised that this information was generated on 2018-01-10 and may be subject to change.
A systematic review on the influence of trial quality on the effect of garlic on blood pressure

S. Simons*, H. Wollersheim, T. Thien

Departments of 1Pulmonary Diseases and 1General Internal Medicine, and 2Scientific Institute on Quality of Healthcare, Radboud University Medical Centre, PO Box 9101, 6500 HB Nijmegen, the Netherlands, *corresponding author: tel.: +31 (0)24-361 45 79, fax: +31 (0)24-361 03 24, e-mail: S.Simons@long.umcn.nl

ABSTRACT

Background: Garlic is a widely used herbal product for hypertension. Previous meta-analyses on the effect of garlic on blood pressure (BP) have been contradictory however. We hypothesised that methodological deficiencies may have contributed to this disagreement. We therefore evaluated whether trials reporting on the effect of garlic on BP had sufficient methodological qualities and a proper description of BP determination.

Methods: MEDLINE, EMBASE, AMED, the COCHRANE library, IBIDS and CINAHL were systematically searched for trials reporting on the effect of garlic on BP. Both the methodological quality and the quality of blood pressure measurement were appraised using predefined quality scores.

Results: 32 studies were identified. Of these studies, 13 were included previously by other meta-analyses. The methodological quality of the studies was poor. Only four trials had adequate allocation concealment, no single trial reported an intention-to-treat analysis and blinding of the evaluators was done in three trials only. Moreover, half of the studies did not report any data on BP measurement. No trials reported on the arm level. Body position was described most often. All trials fulfilling a predefined cutoff point were conducted in normotensive subjects.

Conclusion: The effect of garlic on blood pressure cannot be ascertained. Previous meta-analyses have been based on trials with inadequate study designs, methodological deficiencies and with too little information about blood pressure measurement. In our view, use of garlic cannot be recommended as antihypertensive advice for hypertensive patients in daily practice.

KEYWORDS

Blood pressure, hypertension, garlic, systematic review, quality analysis

INTRODUCTION

For the treatment of hypertension, non-pharmacological therapeutic advice is the initial therapeutic approach in patients with an increased cardiovascular risk. A number of such lifestyle recommendations clearly reduce blood pressure, e.g. weight reduction and regular physical activity. Food supplementation could provide another, less strenuous non-pharmacological intervention. Garlic is the second most used herb taken by patients with cardiovascular disease. Moreover, garlic might lower blood pressure via the conversion of garlic-derived organic polysulphides into hydrogen sulphide by the red blood cell leading to vasorelaxation. It could therefore be a potential target as an antihypertensive food supplement. Indeed, several trials suggest possible short-term effects of garlic on cardiovascular risk factors. The effects of garlic on blood pressure specifically have been summarised in three systematic reviews. However, the results were contradictory. This phenomenon has been evaluated previously by Linde et al. In their analysis on discordant conclusions of systematic reviews in complementary medicine, they concluded that these differences were mostly due to differences in inclusion and exclusion criteria. We wondered whether this could also be explained by the different approaches to appraise trial quality, because trials in complementary medicine have low methodological quality and because these deficiencies may translate into biased findings in systematic reviews.
Besides, none of these reviews looked at the quality of blood pressure measurement. Recently, Wood et al. demonstrated that bias in intervention effects is only seen in trials using subjective outcomes, i.e. physician assessed disease outcome. During blood pressure measurements many errors can occur resulting in varying blood pressures making it amenable to faulty outcomes. Thus when addressing the bias in blood pressure trials introduced by methodological deficiencies, proper blood pressure measurements seem to be an additional quality criterion. We therefore conducted an analysis of the methodological quality of trials reporting on the effect of garlic on blood pressure using multiple strict quality criteria. The present study aims to answer the following four questions:

- Does garlic lower blood pressure in humans?
- What is the methodological quality of human trials measuring the effect of garlic on blood pressure?
- What information about criteria for blood pressure measurements is presented in these garlic studies?
- Did the inclusion of methodologically poor studies affect the conclusions of previous systematic reviews?

**METHODS**

**Literature search**

MEDLINE (1966-2008), EMBASE (1980-2008), CINAHL (1982-2008), IBIDS (2008), AMED (1985-2008) and the COCHRANE library were searched from March 2008 until May 2008 using search strategies mentioned in Table 1. The search was last updated in January 2009 by two reviewers independently. References from garlic reviews and eligible trials were searched for additional articles.

**Selection criteria**

Studies in adults measuring the effect of garlic on blood pressure were considered eligible. Inclusion was limited to studies lasting more than eight weeks and with more than 20 participants. In crossover trials each parallel arm had to last more than eight weeks. Trials were selected independently by two reviewers on the basis of abstracts.

**Data extraction**

Two reviewers identified eligible trials. Data extraction and quality assessment were done by two reviewers. When results of a scoring card (see: quality analysis) deviated more than one point, a third independent reviewer was consulted. Differences were resolved by consensus. In case of different conclusions, consensus could always be reached.

**Quality analysis**

Both the methodological quality and the quality of blood pressure measurement were assessed using two different scoring systems. Methodological quality was assessed using a scoring card derived from the Cochrane checklist ‘the assessment of a randomised trial’. The card consisted of nine items mentioned in Table 2. A trial had to describe each item specifically. Moreover, both the number as well as the reasons for dropping out had to be recorded. Trial arms had to be similar in age, sex, blood pressure and modifying factors (smoking, diabetes, hypercholesterolaemia, overweight, alcohol, cardiovascular comorbidity; at least four mentioned). Each study could score a maximum of nine points.

**METHODS**

**Literature search**

MEDLINE (1966-2008), EMBASE (1980-2008), CINAHL (1982-2008), IBIDS (2008), AMED (1985-2008) and the COCHRANE library were searched from March 2008 until May 2008 using search strategies mentioned in Table 1. The search was last updated in January 2009 by two reviewers independently. References from garlic reviews and eligible trials were searched for additional articles.

**Selection criteria**

Studies in adults measuring the effect of garlic on blood pressure were considered eligible. Inclusion was limited to studies lasting more than eight weeks and with more than 20 participants. In crossover trials each parallel arm had to last more than eight weeks. Trials were selected independently by two reviewers on the basis of abstracts.

**Data extraction**

Two reviewers identified eligible trials. Data extraction and quality assessment were done by two reviewers. When results of a scoring card (see: quality analysis) deviated more than one point, a third independent reviewer was consulted. Differences were resolved by consensus. In case of different conclusions, consensus could always be reached.

**Quality analysis**

Both the methodological quality and the quality of blood pressure measurement were assessed using two different scoring systems. Methodological quality was assessed using a scoring card derived from the Cochrane checklist ‘the assessment of a randomised trial’. The card consisted of nine items mentioned in Table 2. A trial had to describe each item specifically. Moreover, both the number as well as the reasons for dropping out had to be recorded. Trial arms had to be similar in age, sex, blood pressure and modifying factors (smoking, diabetes, hypercholesterolaemia, overweight, alcohol, cardiovascular comorbidity; at least four mentioned). Each study could score a maximum of nine points.

**METHODS**

**Literature search**

MEDLINE (1966-2008), EMBASE (1980-2008), CINAHL (1982-2008), IBIDS (2008), AMED (1985-2008) and the COCHRANE library were searched from March 2008 until May 2008 using search strategies mentioned in Table 1. The search was last updated in January 2009 by two reviewers independently. References from garlic reviews and eligible trials were searched for additional articles.

**Selection criteria**

Studies in adults measuring the effect of garlic on blood pressure were considered eligible. Inclusion was limited to studies lasting more than eight weeks and with more than 20 participants. In crossover trials each parallel arm had to last more than eight weeks. Trials were selected independently by two reviewers on the basis of abstracts.

**Data extraction**

Two reviewers identified eligible trials. Data extraction and quality assessment were done by two reviewers. When results of a scoring card (see: quality analysis) deviated more than one point, a third independent reviewer was consulted. Differences were resolved by consensus. In case of different conclusions, consensus could always be reached.

**Quality analysis**

Both the methodological quality and the quality of blood pressure measurement were assessed using two different scoring systems. Methodological quality was assessed using a scoring card derived from the Cochrane checklist ‘the assessment of a randomised trial’. The card consisted of nine items mentioned in Table 2. A trial had to describe each item specifically. Moreover, both the number as well as the reasons for dropping out had to be recorded. Trial arms had to be similar in age, sex, blood pressure and modifying factors (smoking, diabetes, hypercholesterolaemia, overweight, alcohol, cardiovascular comorbidity; at least four mentioned). Each study could score a maximum of nine points.

**METHODS**

**Literature search**

MEDLINE (1966-2008), EMBASE (1980-2008), CINAHL (1982-2008), IBIDS (2008), AMED (1985-2008) and the COCHRANE library were searched from March 2008 until May 2008 using search strategies mentioned in Table 1. The search was last updated in January 2009 by two reviewers independently. References from garlic reviews and eligible trials were searched for additional articles.

**Selection criteria**

Studies in adults measuring the effect of garlic on blood pressure were considered eligible. Inclusion was limited to studies lasting more than eight weeks and with more than 20 participants. In crossover trials each parallel arm had to last more than eight weeks. Trials were selected independently by two reviewers on the basis of abstracts.

**Data extraction**

Two reviewers identified eligible trials. Data extraction and quality assessment were done by two reviewers. When results of a scoring card (see: quality analysis) deviated more than one point, a third independent reviewer was consulted. Differences were resolved by consensus. In case of different conclusions, consensus could always be reached.

**Quality analysis**

Both the methodological quality and the quality of blood pressure measurement were assessed using two different scoring systems. Methodological quality was assessed using a scoring card derived from the Cochrane checklist ‘the assessment of a randomised trial’. The card consisted of nine items mentioned in Table 2. A trial had to describe each item specifically. Moreover, both the number as well as the reasons for dropping out had to be recorded. Trial arms had to be similar in age, sex, blood pressure and modifying factors (smoking, diabetes, hypercholesterolaemia, overweight, alcohol, cardiovascular comorbidity; at least four mentioned). Each study could score a maximum of nine points.

**METHODS**

**Literature search**

MEDLINE (1966-2008), EMBASE (1980-2008), CINAHL (1982-2008), IBIDS (2008), AMED (1985-2008) and the COCHRANE library were searched from March 2008 until May 2008 using search strategies mentioned in Table 1. The search was last updated in January 2009 by two reviewers independently. References from garlic reviews and eligible trials were searched for additional articles.

**Selection criteria**

Studies in adults measuring the effect of garlic on blood pressure were considered eligible. Inclusion was limited to studies lasting more than eight weeks and with more than 20 participants. In crossover trials each parallel arm had to last more than eight weeks. Trials were selected independently by two reviewers on the basis of abstracts.

**Data extraction**

Two reviewers identified eligible trials. Data extraction and quality assessment were done by two reviewers. When results of a scoring card (see: quality analysis) deviated more than one point, a third independent reviewer was consulted. Differences were resolved by consensus. In case of different conclusions, consensus could always be reached.

**Quality analysis**

Both the methodological quality and the quality of blood pressure measurement were assessed using two different scoring systems. Methodological quality was assessed using a scoring card derived from the Cochrane checklist ‘the assessment of a randomised trial’. The card consisted of nine items mentioned in Table 2. A trial had to describe each item specifically. Moreover, both the number as well as the reasons for dropping out had to be recorded. Trial arms had to be similar in age, sex, blood pressure and modifying factors (smoking, diabetes, hypercholesterolaemia, overweight, alcohol, cardiovascular comorbidity; at least four mentioned). Each study could score a maximum of nine points.
Roche et al. have reported on the quality of the blood pressure measurement in medical literature. During a preliminary data analysis, however, it became clear that no single study sufficiently fulfilled the criteria used in that study. Therefore, from the list used by Roche et al., five essential criteria were formulated based on a priorisation procedure (table 3). The methodology section of each trial was screened for the description of blood pressure measurement by two reviewers independently. A study could score a maximum of five points if each criterion was written specifically.

Roche et al. have reported on the quality of the blood pressure measurement in medical literature. During a preliminary data analysis, however, it became clear that no single study sufficiently fulfilled the criteria used in that study. Therefore, from the list used by Roche et al., five essential criteria were formulated based on a priorisation procedure (table 3). The methodology section of each trial was screened for the description of blood pressure measurement by two reviewers independently. A study could score a maximum of five points if each criterion was written specifically.

<table>
<thead>
<tr>
<th>Quality criterion</th>
<th>Number of studies fulfilling criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure device and accuracy mentioned</td>
<td>6</td>
</tr>
<tr>
<td>Body position reported</td>
<td>13</td>
</tr>
<tr>
<td>Measurement with arm at heart level</td>
<td>0</td>
</tr>
<tr>
<td>Rest period before measurement</td>
<td>7</td>
</tr>
<tr>
<td>Number of readings reported</td>
<td>9</td>
</tr>
</tbody>
</table>

Comparison with earlier systematic reviews
We hypothesised that conclusions drawn by previous authors were based on differences in inclusion and exclusion criteria due to differences in evaluating trial quality. Methodological weaknesses were identified in the two earlier meta-analyses by examining if the trials included by these authors lacked our quality criteria. Because one previous review did not use a quality assessment this review was not included in the analysis.

RESULTS

Figure 1 shows the selection process. Of the 203 trials found through the initial database search, 165 possible garlic trials were identified. The crosschecking of references and reviews revealed another 17 trials. A total of 137 studies did not meet the inclusion criteria. Another 13 trials were excluded for various reasons mentioned in figure 1. The remaining 32 trials were further analysed.

Study description
Table 4 describes the trials reporting on the effect of garlic on blood pressure. Sample size ranged from 23 to 862 subjects and trials lasted from 8 to 156 weeks. Fifteen trials had blood pressure as a primary objective and 16 trials comprised of hypertensive patients. Different study designs were used: 21 double-blind, randomised, placebo-controlled, parallel trials; two randomised placebo-controlled crossover trials; one single-blind, placebo-controlled trial; one double-blind, randomised trial against an active component; six uncontrolled before-after trials; one controlled before-after trial. Only one trial was a priori set up in a randomised placebo-controlled fashion to test solely whether garlic lowers blood pressure in a hypertensive population.

Some trials lacked important data. Fourteen trials did not report any numerical blood pressure data. Two studies reported no demographic data. 14 studies did not explicitly exclude patients on hypertensive medication and in two trials such medication was allowed.

Table 4. Summary of the 32 clinical trials of more than eight weeks examining the effect of garlic ingestion on blood pressure

<table>
<thead>
<tr>
<th>Author (publication year) [reference number]</th>
<th>Sample size</th>
<th>Primary objective</th>
<th>Blood pressure</th>
<th>Intervention (total daily intake in grams)</th>
<th>Control group (total daily intake in grams)</th>
<th>Co-interventions</th>
<th>Trial length (in weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler et al. (1997)15</td>
<td>50</td>
<td>LP</td>
<td>N</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>Primrose oil</td>
<td>12</td>
</tr>
<tr>
<td>Auer et al. (1990)16</td>
<td>47</td>
<td>LP + BP</td>
<td>H</td>
<td>GP (0.6)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Bordia et al. (1989)17</td>
<td>42</td>
<td></td>
<td>NND</td>
<td>GO (NND)</td>
<td>PL</td>
<td>NG</td>
<td>15</td>
</tr>
<tr>
<td>Brewitt et al. (1991)18</td>
<td>862</td>
<td></td>
<td>H</td>
<td>GP (6.12)</td>
<td>No placebo group</td>
<td>NG</td>
<td>26</td>
</tr>
<tr>
<td>Cheng et al. (2006)19</td>
<td>79</td>
<td>LP + BP</td>
<td>H</td>
<td>GP (0.012) + diuretic</td>
<td>Diuretic</td>
<td>Diuretic</td>
<td>52</td>
</tr>
<tr>
<td>Czerny et al. (1996)20</td>
<td>100</td>
<td>LP</td>
<td>H</td>
<td>GO (0.4) + lecithin</td>
<td>PL</td>
<td>NG</td>
<td>14</td>
</tr>
<tr>
<td>De A Santos et al. (1993)21</td>
<td>60</td>
<td>LP + BP</td>
<td>N + H</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>DA</td>
<td>24</td>
</tr>
<tr>
<td>De A Santos et al. (1995)22</td>
<td>80</td>
<td>LP + BP</td>
<td>N + H</td>
<td>GP (1.8)</td>
<td>GO (0.006)</td>
<td>NG</td>
<td>16</td>
</tr>
<tr>
<td>Dhawan et al. (2004)23</td>
<td>40</td>
<td>LP + BP</td>
<td>N + H</td>
<td>Garlic pearls (0.4)</td>
<td>Garlic pearls (0.4)</td>
<td>NG</td>
<td>8</td>
</tr>
<tr>
<td>Durak et al. (2003)24</td>
<td>23</td>
<td>LP</td>
<td>N + H</td>
<td>Garlic extract (10)</td>
<td>No control group</td>
<td>NG</td>
<td>16</td>
</tr>
<tr>
<td>Gardner et al. (2001)25</td>
<td>33</td>
<td>LP</td>
<td>N</td>
<td>GP (0.3-1.0)</td>
<td>PL</td>
<td>DA</td>
<td>12</td>
</tr>
<tr>
<td>Grünwald et al. (1992)26</td>
<td>48</td>
<td>LP + BP</td>
<td>N + H</td>
<td>GP (0.6)</td>
<td>No placebo group</td>
<td>NG</td>
<td>18</td>
</tr>
<tr>
<td>Holzgartner et al. (1993)27</td>
<td>98</td>
<td>LP</td>
<td>N + H</td>
<td>GP (0.9)</td>
<td>Bezafibrate (0.6)</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Isaacs et al. (1998)28</td>
<td>30</td>
<td>LP</td>
<td>N</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>DA</td>
<td>12</td>
</tr>
<tr>
<td>Jabbari et al. (2005)29</td>
<td>50</td>
<td>LP</td>
<td>N</td>
<td>Raw garlic (1)</td>
<td>Raw garlic (1)</td>
<td>NG</td>
<td>8</td>
</tr>
<tr>
<td>Jain et al. (1993)30</td>
<td>42</td>
<td>LP + BP</td>
<td>N</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Kandziora (1988)31</td>
<td>40</td>
<td>BP</td>
<td>H</td>
<td>GP (0.6) + diuretic</td>
<td>Diuretic-reserpin combination</td>
<td>Lifestyle advice</td>
<td>12</td>
</tr>
<tr>
<td>Kiesewetter et al. (1993)32</td>
<td>80</td>
<td></td>
<td>NND</td>
<td>GO (0.8)</td>
<td>PL + diuretic</td>
<td>Low salt diet</td>
<td>12</td>
</tr>
<tr>
<td>Lutomski (1984)33</td>
<td>102</td>
<td>NG</td>
<td>N + H</td>
<td>GP (0.3) + rutin</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Macan et al. (2006)34</td>
<td>52</td>
<td>Warfarin safety</td>
<td>N + H</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Mansell et al. (1996)35</td>
<td>60</td>
<td>BP</td>
<td>NND</td>
<td>GO (0.9)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>McManus et al. (1994)36</td>
<td>42</td>
<td>BP</td>
<td>NND</td>
<td>GO (0.9)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Moziwicki et al. (1988)37</td>
<td>NND</td>
<td>NG</td>
<td>NND</td>
<td>GO (0.05-0.18) or GO (3)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Sarath et al. (1994)38</td>
<td>72</td>
<td>LP</td>
<td>N</td>
<td>GP (0.6)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Simons et al. (1995)39</td>
<td>31</td>
<td>LP</td>
<td>N</td>
<td>GP (0.9)</td>
<td>Lactose</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Steiner et al. (1996)40</td>
<td>52</td>
<td>LP + BP</td>
<td>N + H</td>
<td>GP (2.4)</td>
<td>PL</td>
<td>DA</td>
<td>24</td>
</tr>
<tr>
<td>Superko et al. (2000)41</td>
<td>50</td>
<td>LP</td>
<td>N</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>DA</td>
<td>12</td>
</tr>
<tr>
<td>Turner et al. (2004)42</td>
<td>75</td>
<td>LP + BP</td>
<td>N</td>
<td>GP (0.92)</td>
<td>PL</td>
<td>NG</td>
<td>12</td>
</tr>
<tr>
<td>Vorberg et al. (1990)43</td>
<td>40</td>
<td>NG</td>
<td>N + H</td>
<td>GP (0.9)</td>
<td>PL</td>
<td>NG</td>
<td>16</td>
</tr>
<tr>
<td>Zhang et al. (2000)45</td>
<td>36</td>
<td>BP</td>
<td>N</td>
<td>GO (0.012)</td>
<td>PL</td>
<td>NG</td>
<td>16</td>
</tr>
<tr>
<td>Ziaei et al. (2001)46</td>
<td>100</td>
<td>LP</td>
<td>N</td>
<td>GP (0.8)</td>
<td>PL</td>
<td>NG</td>
<td>8</td>
</tr>
</tbody>
</table>

BP = blood pressure; DA = dietary advice with adequate instruction; GO = garlic oil; GP = garlic powder; H = hypertensive; LP = lipids; N = normotensive; NND = no numerical data available; NG = data not given; PL = placebo. *Crossover studies; only the first parallel arm.

Methodological quality

Figure 2 shows how well studies performed on the methodology scale. No trials scored maximum. The median score was four points (range 0-8). Eleven trials scored five or more points.21,25,27-28,30,33,35,40,43-45 Trials scored low for intention-to-treat analysis, allocation concealment (i.e., shielding those who admit participants to a study from knowing the upcoming assignments) and for the blinding of evaluators. This is shown in table 2. Randomisation21,25,27-28,30,33,40,41,43,45 and providing the same treatment among treatment arms21,25,28,30,41,43,45 were fairly well described in the selected studies.

Blood pressure measurement

Figure 3 depicts the spread of scores on reporting the five blood pressure criteria. Scoring was poor. Half of the trials did not report any criterion17,21,24,26-27,29,34,35,41-45 and only six trials scored three points or more.23,25,30,40,43,45 A subdivision of the scoring system is shown in table 3. Information about the five criteria was scarce; no trial reported arm height during measurement and less than a quarter described the blood pressure device25,30,40,41,43,45 or the resting period.23,25,30,33,40,43,45 Body position was most often cited.21,25,26,28,29,33,35,40,41,43,44
Comparison with earlier meta-analysis
Of the 32 studies, 13 have been studied by either Ried et al. or Silagy and Neil. The latter also included one study lasting less than eight weeks. The quality of these studies is summarised in table 5. All trials lack proper allocation concealment, blinding of evaluators and intention-to-treat analysis. Moreover, there was little information about the quality of blood pressure measurement. Only two trials mentioned resting period, the device used, body position of the patient and number of readings. Trials with highest scores in both our quality assessments are summarised in table 6. For clarity cutoff points of 5 and 3 were used for the methodology score and the blood pressure score respectively. Garlic did not lower blood pressure in any selected trial. However, none of these trials were performed in hypertensive subjects.

DISCUSSION
Our main conclusion is that the hypotensive effect of garlic cannot be ascertained, because conclusions from previous meta-analyses have been based on trials with inadequate study designs, with methodological defects and with insufficient information on blood pressure measurement. Moreover, trials with the best methodology were performed in normotensive subjects.

This study shows that designs of trials on the effect of garlic on blood pressure have important flaws. Few trials used only hypertensive subjects. Only one trial was set up in a randomised, placebo-controlled, double-blind fashion to evaluate the effect of garlic on blood pressure in a hypertensive population. Although other study designs may be used, these are prone to the introduction of bias. Moreover, it may be questionable, but not entirely ruled out, that effects found in normotensive subjects can be extrapolated to hypertensive patients.

Secondly, the methodological quality was poor. No trials reported intention-to-treat analysis and only a minority had an adequate allocation concealment and blinding of evaluators. This is in agreement with two previous systematic reviews on the effect of garlic on blood pressure. Using another methodology scale, Silagy and Neil scored for randomisation, intention-to-treat analysis and for blinding of the evaluators. None of their eight included trials scored positive on all points and they therefore concluded that ‘quality assessment of the trials was generally poor’. These trials also scored weak in our rating scales. Another systematic review by Ackermann et al. did not perform a systematic quality analysis. However, they do provide some overview of trial quality mainly corroborating with our findings. Ackermann et al. refrained from carrying out a meta-analysis because ‘about half of the studies did not present numerical data, multiple blood pressure measurements were used and few studies had a priori hypotheses related to blood pressure.’

In trials on complementary medicine in general, others have also reported low methodological quality. Linde et al. reviewed trials on homeopathy, acupuncture and herbal medicine, garlic not included. They concluded that most trials had inadequate allocation concealment, poor randomisation procedures and gave little information about dropouts. In a more comprehensive review, Gagnier et al. analysed 1321 English trials on herbal medicine, garlic included. They concluded that trials in herbal medicine provided less than half of the necessary information in their reports. Of the ten herbs studied, garlic trials scored second worst. Allocation concealment, randomisation and
blinding of the evaluators, as in the present study, were poorly described.

Our conclusion on the methodological quality contrasts with the work by Ried et al., however. In their systematic review they also used guidelines by the Cochrane Collaboration for assessing trial quality. They concluded on the basis of their analysis that trial quality of their included trials was ‘generally high’. However, they only provided information about blinding, randomisation and blood pressure as a primary outcome. Our analysis shows that what trials mostly lack is an adequate allocation concealment, blinding of the evaluators and proper use of an intention-to-treat analysis. It has been shown that inadequate allocation concealment or blinding may lead to exaggerated treatment effects. When taking into account all nine quality criteria, the performance of the included garlic trials is, at least, equivocal.

The third conclusion that can be drawn from our results is that trials provide insufficient information about the technique used to measure blood pressure. The absence of proper information about blood pressure measurements is worrying since factors such as arm position may influence blood pressure measurements by up to 10 mmHg. The fact that only 15 studies had blood pressure as a primary outcome might explain this absence. One could argue it is not the blood pressure per se that counts but rather the difference between the measurements. However, incorrect

### Table 5. Quality analysis of the 14 included trials in the meta-analyses performed by Ried et al. or by Silagy and Neil

<table>
<thead>
<tr>
<th>Author (year) [reference]</th>
<th>Included by Ried (R) or Silagy (S)</th>
<th>Total points quality scale (max. 9)</th>
<th>Lacking items</th>
<th>Total points blood pressure scale (max. 5)</th>
<th>Lacking items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auer (1990)</td>
<td>R + S</td>
<td>4</td>
<td>AC, EB, CG, D, IT</td>
<td>1</td>
<td>DM, AH, RP, NR</td>
</tr>
<tr>
<td>De A Santos (1993)</td>
<td>S</td>
<td>6</td>
<td>AC, EB, IT</td>
<td>0</td>
<td>DM, AH, BP, RP, NR,</td>
</tr>
<tr>
<td>Jain (1993)</td>
<td>R + S</td>
<td>5</td>
<td>AC, EB, D, IT</td>
<td>4</td>
<td>AH</td>
</tr>
<tr>
<td>Kiesewetter (1991)</td>
<td>S</td>
<td>4</td>
<td>AC, EB, CG, D, IT</td>
<td>0</td>
<td>DM, AH, BP, RP, NR</td>
</tr>
<tr>
<td>Kiesewetter (1993)</td>
<td>R</td>
<td>5</td>
<td>AC, EB, D, IT</td>
<td>2</td>
<td>DM, AH, NR</td>
</tr>
<tr>
<td>Saradeth (1996)</td>
<td>R</td>
<td>4</td>
<td>AC, EB, CG, D, IT</td>
<td>0</td>
<td>DM, AH, BP, RP, NR</td>
</tr>
<tr>
<td>Simons (1999)</td>
<td>R</td>
<td>5</td>
<td>AC, EB, CG, IT</td>
<td>4</td>
<td>AH</td>
</tr>
<tr>
<td>Steiner (1996)</td>
<td>R</td>
<td>4</td>
<td>AC, EB, CG, D, IT</td>
<td>1</td>
<td>BP, AH, RP, NR</td>
</tr>
<tr>
<td>Vorberg (1990)</td>
<td>R + S</td>
<td>5</td>
<td>AC, EB, CG, IT</td>
<td>1</td>
<td>DM, AH, RP, NR</td>
</tr>
<tr>
<td>Zhang (2000)</td>
<td>R</td>
<td>6</td>
<td>AC, EB, D, IT</td>
<td>3</td>
<td>BP, AH</td>
</tr>
</tbody>
</table>

AC = allocation concealment; PB = patient blinding; EB = evaluators blinding; D = dropouts; CG = comparable groups; IT = intention-to-treat analysis; ST = same treatment of groups; DM = device mentioned; AH = arm at heart level; BP = body position; RP = resting period; NR = number of reading. Not included in our systematic review because of a treatment period of less than eight weeks.

### Table 6. Summary of the five clinical trials on the effect of garlic on blood pressure with the highest methodological quality score and highest score for the quality in reporting blood pressure measurements

<table>
<thead>
<tr>
<th>Author (year) [reference]</th>
<th>Blood pressure a primary goal</th>
<th>Hypertensive population</th>
<th>Methodological quality*</th>
<th>Blood pressure quality**</th>
<th>Garlic effective in lowering blood pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gardner (2001)</td>
<td>No</td>
<td>No</td>
<td>7</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Jain (1993)</td>
<td>No</td>
<td>No</td>
<td>5</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Simons (1999)</td>
<td>No</td>
<td>No</td>
<td>5</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Turner (2004)</td>
<td>Yes</td>
<td>No</td>
<td>8</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Zhang (2000)</td>
<td>Yes</td>
<td>No</td>
<td>6</td>
<td>3</td>
<td>No</td>
</tr>
</tbody>
</table>

*Score based on the criteria proposed by the Dutch Cochrane society. No points indicates a high suspicion of bias. A cutoff value of five points or more was chosen. **Blood pressure quality was scored on reporting on exact device, body position, arm at heart level, resting period and number of readings reported. A maximum of five points could be obtained. A cutoff value of three points was chosen.
blood pressure measurements would inherently lead to
data pollution of normotensive and hypertensive subjects
making conclusions impossible. Besides, none of the
studies included in the present review described that blood
pressure was measured similarly every time.

To our knowledge, the present study is the first review
that systematically assessed blood pressure measurement
in garlic trials. Of the three previous systematic reviews
on garlic and blood pressure, only Ried et al. provide
data about blood pressure measurements, but they do not
draw the conclusion that the information was insufficient.
Roche et al. have reviewed the reporting of blood pressure
measurements in leading English medical journals. In
their analysis of 116 papers, device accuracy and validation
or reporting of arm level was also poorly described;
device accuracy was only reported in 3% and arm level
in 5%. As in the present study, both body position and
number of readings were best reported. Also in a review
of Brazilian medical literature, Holanda et al. showed
articles lacked important data; in only half of the studies
the type of sphygmomanometer or the number of readings
were mentioned and only a quarter described the body
position.

Did the inclusion of these garlic trials affect the conclusions
of previous meta-analyses? In our view the absence of a
proper methodology may have indeed affected outcome.
Both Schulz et al. and Moher et al. have shown that
inadequate allocation concealment and lack of blinding
may lead to exaggerated results. In the present study it
was shown that garlic trials scored poorly in both criteria.
Biased results in trials may also affect results of systematic
reviews. According to the criteria put forward in the
present study, only five trials provide sufficient quality data
(table 6). These studies do not show an effect of
garlic on blood pressure. We refrained from conducting a
meta-analysis, however, because all trials were performed
on normotensive subjects. In our view these trials are not
suitable to answer the only clinically relevant question
whether garlic lowers blood pressure in hypertensive
patients.

Our analysis has several limitations. Our scoring systems
inherently had some subjective elements. We minimised
this effect by using two independent evaluators and a
standardised checklist. Second, the absence of reporting
procedures in the garlic trials does not necessarily imply
these procedures were not done. Our criteria might
have been biased towards precisely written studies.
It is our opinion, though, that transparency is a key
element in conducting trials. In such a way, clear writing
might in itself be a quality criterion. Deficient reporting
generally embodies imperfect methodologies. Thirdly,
the applicability of scores to appraise the methodological
quality has been challenged by others, arguing that
high scores do not necessarily represent valid trials. They
propose that relevant methodological aspects should be
assessed individually. We therefore used both, because
in our view an overall composite score nevertheless gives a
pragmatic and visual tool for a global quality assessment.

CONCLUSION

It is our view that garlic cannot be recommended to lower
blood pressure in the daily practice of physicians working
in the field of hypertension treatment, given the low
methodological quality, the lack of information about blood
pressure measurement and the absence of methodological
sound trials in hypertensive patients.

ACKNOWLEDGEMENTS

We would like to thank Professor J. Lenders and Dr J.
Deinum for their assistance in prioritisising quality criteria
for blood pressure measurement.

REFERENCES


