A systematic review of randomized controlled trials on curative and health enhancement effects of forest therapy

Hiroharu Kamioka1
Kiichiro Tsutani2
Yoshihito Mutoh3
Takuya Honda4
Nobuyoshi Shiozawa5
Shinpei Okada6
Sang-Jun Park6
Jun Kitayuguchi7
Masamitsu Kamada8
Hiroyasu Okuizumi9
Shuichi Handa9

1Faculty of Regional Environment Science, Tokyo University of Agriculture, Tokyo, 2Department of Drug Policy and Management, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo, 3Todai Policy Alternatives Research Institute, The University of Tokyo, Tokyo, 4Japanese Society for the Promotion of Science, Tokyo, 5Food Labeling Division, Consumer Affairs Agency, Cabinet Office, Government of Japan, Tokyo, 6Physical Education and Medicine Research Foundation, Nagano, 7Physical Education and Medicine Research Center Unnan, Shimane, 8Department of Environmental and Preventive Medicine, Shimane University School of Medicine, Shimane, 9Mimaki Onsen (Spa) Clinic, Tomi City, Nagano, Japan

Objective: To summarize the evidence for curative and health enhancement effects through forest therapy and to assess the quality of studies based on a review of randomized controlled trials (RCTs).

Study design: A systematic review based on RCTs.

Methods: Studies were eligible if they were RCTs. Studies included one treatment group in which forest therapy was applied. The following databases – from 1990 to November 9, 2010 – were searched: MEDLINE via PubMed, CINAHL, Web of Science, and Ichushi-Web. All Cochrane databases and Campbell Systematic Reviews were also searched up to November 9, 2010.

Results: Two trials met all inclusion criteria. No specific diseases were evaluated, and both studies reported significant effectiveness in one or more outcomes for health enhancement. However, the results of evaluations with the CONSORT (Consolidated Standards of Reporting Trials) 2010 and CLEAR NPT (A Checklist to Evaluate a Report of a Nonpharmacological Trial) checklists generally showed a remarkable lack of description in the studies. Furthermore, there was a problem of heterogeneity, thus a meta-analysis was unable to be performed.

Conclusion: Because there was insufficient evidence on forest therapy due to poor methodological and reporting quality and heterogeneity of RCTs, it was not possible to offer any conclusions about the effects of this intervention. However, it was possible to identify problems with current RCTs of forest therapy, and to propose a strategy for strengthening study quality and stressing the importance of study feasibility and original check items based on characteristics of forest therapy as a future research agenda.

Keywords: forest therapy, randomized controlled trial, curative effect, health enhancement

Introduction

Over the years, recreation activity and relaxation in a forest environment called “forest therapy” or “shinrin-yoku” (forest-air bathing and forest-landscape watching/walking) have become a kind of climatherapy or nature therapy, and are popular methods for many urban people with mental stress conditions. The fields of preventive and alternative medicine have also shown an interest in the therapeutic effects of forest therapy.1

The green landscape may help one recovery from stress by lowering blood pressure,2,3 increasing alpha brain wave amplitude,2 and reducing muscle tension.2 Exposure to negative ions may also enhance parasympathetic nervous activity and decrease blood glucose levels.4,5 A study reported that forest environments may contribute to the maintenance of health and wellbeing (eg, by reducing hostility and
depression which are risk factors for coronary heart disease, or by improving overall emotions, particularly among populations with poor mental health). In addition, a recent study reported that forest bathing trips increase natural killer (NK) activity, which was mediated by an increase in the number of NK cells and the levels of intracellular anticancer proteins and phytoncides released from trees. The decreased production of stress hormones may also partially contribute to the increased NK activity.

It is well known in research design that evidence grading is highest for a systematic review with meta-analysis of randomized controlled trials (RCTs). Although many studies have reported the effects of forest therapy, there is no systematic review of the evidence based on RCTs. The objective of this review was to summarize the evidence from RCTs on the curative and health enhancement effects of forest therapy, and to assess the quality of those trials.

Methods
Criteria for considering studies included in this review
Types of studies
Studies were eligible if they were RCTs.

Types of participants
There was no restriction on participants (patients or healthy participants).

Types of intervention and language
Studies included at least one treatment group in which forest therapy was applied. Any kind of forest therapy (not only taking in the forest atmosphere and forest bathing but also experiencing artificial roof gardens, dwarfed tree appreciation, and green-scene simulations) was permitted and defined as intervention. The use of medication, alternative therapies, or lifestyle changes described were comparable in the groups studied. Therefore, no restriction on the basis of language.

Types of outcome measures
The primary outcome for measurement of effectiveness was the benefit of forest therapy. Beneficial outcome measures included the following: blood pressure, alpha brain wave amplitude, parasympathetic nervous activity, muscle tension, blood glucose level, NK activity, pain, and mental health status (mood, self-esteem, emotion, and subjective wellbeing or happiness).

Search methods for identification of studies
Bibliographic database
The following databases – from 1990 to November 9, 2010 – were searched: MEDLINE via PubMed, CINAHL, Web of Science, and Ichushi-Web (in Japanese). The International Committee of Medical Journal Editors recommended uniform requirements for manuscripts submitted to biomedical journals in 1993. Articles published after 1990 were selected because it appeared that the International Committee of Medical Journal Editors’ recommendation had been adopted by the relevant researchers, thus strengthening the quality of reports. The Cochrane Database of Systematic Reviews (Cochrane Reviews), the Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials (CENTRAL), the Cochrane Methodology Register, the Health Technology Assessment Database, the National Health Service Economic Evaluation Database, and Campbell Systematic Reviews (the Campbell Collaboration) were also searched up to November 9, 2010.

All searches were performed by a specific researcher (hospital librarian) who was qualified in medical information handling, and who was sophisticated in searches of clinical trials.

Search strategies
The special search strategies contained the following elements and terms for MEDLINE, CINAHL, Web of Science, and Ichushi-Web databases:

I. Search “Nature” [Mesh] and “Baths” [Mesh]
II. Search (Relaxation Therapy [MH] or Relaxation [MH]) and Trees [MH]
III. Search Leisure Activities [MH] and Trees [MH]
IV. Search shinrin-yoku or “forest bathing” or shinrin-yoku
V. Search I or II or III or IV Limits: Publication Date from January 1, 1990 to 2010.

Only keywords about intervention were used for the searches. First, titles and abstracts of identified published articles were reviewed in order to determine the relevance of the articles. Next, references in relevant studies and identified RCTs were screened.

Registry checking (included protocol)
The International Clinical Trials Registry Platform, ClinicalTrials.gov, the University Hospital Medical Information Network Clinical Trial Registry, the Japan Pharmaceutical Information Center Clinical Trials Information,
and the Japan Medical Association Center of Clinical Trials were searched up to November 9, 2010. The International Clinical Trials Registry Platform in the World Health Organization Registry Network meets specific criteria for content, quality and validity, accessibility, unique identification, technical capacity, and administration. Primary registries meet the requirements of the International Committee of Medical Journal Editors. ClinicalTrials.gov is a registry of federally and privately supported clinical trials conducted in the United States and around the world. The University Hospital Medical Information Network Clinical Trial Registry, Japan Pharmaceutical Information Center Clinical Trials Information, and Japan Medical Association Center of Clinical Trials are registries of clinical trials conducted in Japan and around the world.

Reference checking, handsearching, and other
The references of included studies were not checked for further relevant literature. Abstracts published in forest therapies and relevant journals were not handsearched. The National Institutes of Health and the National Center for Complementary and Alternative Medicine in the United States were contacted to inquire about research on forest therapy.

Review methods
Selection of trials
In order to make the final selection of studies for the review, all criteria were applied independently by two authors (NS and JK) to the full text of articles that had passed the first eligibility screening (Figure 1). Disagreements and uncertainties were resolved by discussion with other authors (TH and HK).

Studies were selected when (a) the design was an RCT and (b) one of the interventions was a form of forest therapy. Curative or health enhancement effects were used as a primary outcome measure. Trials that were excluded are presented with reasons for exclusion (Appendix).

Quality assessment of included studies
In order to ensure that variation was not caused by systematic errors in the study design or execution, three review authors (MK, SP, and TH) independently assessed the quality of articles. The quality assessment of these papers was made using the “CONSORT (Consolidated Standards of Reporting Trials) 2010” checklist and the “CLEAR NPT (A Checklist to Evaluate a Report of a Nonpharmacological Trial)” checklist, developed to assess the methodological quality of RCTs and nonpharmacological trials, respectively. Disagreements and uncertainties were resolved by discussion with other authors (HO, SO, and HK).

A meta-analysis could not be performed as the main outcome measures were different and could not be compared between the eligible papers.

Summary of studies and data extraction
Three review authors (MK, SP, and TH) selected the summary from each of the structured abstracts.

Benefit, harm, and withdrawals
The GRADE (Grading of Recommendations Assessment, Development, and Evaluation) Working Group reported that the balance between benefit and harm, quality of evidence, applicability, and the certainty of the baseline risk...
were all considered in judgments about the strength of recommendations. Adverse events, withdrawals, and cost for intervention were especially important information for researchers and users of clinical practice guidelines, and this information was presented with the description of each article.

Analysis
Preplanned stratified analyses were: (a) trials comparing forest therapy with no treatment or waiting list controls, (b) trials comparing different types of forest therapy, and (c) trials comparing forest therapy with other treatment(s) (eg, sea viewing, urban viewing, or living). The results of each RCT were expressed, when possible, as relative risk with corresponding 95% confidence intervals for dichotomous data, and as standardized or weighted mean differences with 95% confidence intervals for continuous data. However, heterogeneous results of studies that were provided by inclusion criteria were not combined.

Results
Study characteristics
The literature searches included 138 potentially relevant articles (Figure 1). Abstracts from those articles were assessed and eleven papers were retrieved for further evaluation (checks for relevant literature). Nine publications were excluded because they did not meet the eligibility criteria (Appendix). Only two studies met all inclusion criteria (Table 1). The language of both of the eligible publications was English. The study participants were all healthy people. Although both studies reported significant effectiveness in their outcomes, these were short duration experiments and the effects could not determined over the long term.

Pretty et al concluded that green exercise has important public and environmental health consequence. Five groups of 20 subjects were exposed to a sequence of 30 scenes projected on a wall whilst exercising on a treadmill. Four categories of scenes were tested: rural pleasant, rural unpleasant, urban pleasant, and urban unpleasant. The control was running without exposure to images. Blood pressure and two psychological measures (Rosenberg Self-Esteem Scale and Profile of Mood States questionnaire) were measured before and after the intervention. Exercise alone significantly reduced blood pressure, increased self-esteem, and had a significant positive effect on four of six mood measures. Both rural and urban pleasant scenes produced a significantly greater positive effect on self-esteem than the exercise-only control, demonstrating a synergistic effect of green exercise in both rural and urban environments. By contrast, both rural and urban unpleasant scenes reduced the positive effect of exercise on self-esteem. The rural unpleasant scenes had the most dramatic effect, depressing the beneficial effects of exercise on three different measures of mood.

Hartig et al compared psychophysiological stress recovery and directed attention restoration in natural and urban field settings using repeated measures of ambulatory blood pressure, emotion, and attention. They reported that walking in a nature reserve initially fostered blood pressure change that indicated greater stress reduction than that afforded by walking in urban surroundings. The positive effect increased and anger decreased in the nature reserve by the end of the walk; the opposite pattern emerged in the urban environment.

Meta-analysis
A meta-analysis could not be performed due to the heterogeneity of the RCTs.

Withdrawals and adverse events
There were no withdrawals (dropouts) reported in either study. A description of adverse events was not presented (Table 1). It was not able to be determined whether or not there were adverse events in the studies.

Costs of intervention
Neither study provided information on the costs of intervention (Table 1).

Quality assessment
Twenty-five items from the CONSORT 2010 checklist were evaluated in more detail (Table 2). This assessment evaluated the quality of how the main findings of the study were summarized in the written report. A lack of description was remarkable for the studies in general. The items for which the description was lacking in both studies were as follows: “identification as a randomized trial in the title;” “structured summary of trial design, methods, results, and conclusions;” “important changes to methods after trial commencement;” “how sample size was determined;” “when applicable, explanation of any interim analyses and stopping guidelines;” “method used to generate the random allocation sequence;” “type of randomization;” “mechanism used to implement the random allocation sequence, describing any steps taken to conceal the sequence until interventions were assigned;” “dates defining the periods of recruitment and follow-up;”
Table 1  Brief summary of articles based on structured abstracts and additional elements

<table>
<thead>
<tr>
<th>Reference</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aim/objective</td>
<td>To examine whether there is synergistic benefit in adopting physical activities whilst being directly exposed to nature</td>
<td>To compare psychophysiological stress recovery and directed attention restoration in natural and urban field settings</td>
</tr>
<tr>
<td>Setting/place</td>
<td>A laboratory at the Department of Biological Sciences, University of Essex (Colchester, United Kingdom)</td>
<td>After the pretreatment phase (a laboratory at the University of California, Irvine, CA), participants went into the environmental phase, where natural and urban environments exist. The natural environment was the Audubon Society’s Starr Ranch Sanctuary (45 minutes from the university by car). Operations were run out of plainly furnished room with windows through which participants could look out onto tree and vegetated hillsides. The urban setting was an area of medium-density professional office and retail development in the City of Orange (45 minutes from the university by car). Operations were run out of quiet, undecorated classrooms without window views</td>
</tr>
<tr>
<td>Participants</td>
<td>100 undergraduates and employees (24.6 ± 0.99 years; 55 females, 45 males) from the university. They were categorized into five conditions: rural pleasant image, rural unpleasant image, urban pleasant image, urban unpleasant image, and no image (control)</td>
<td>112 normotensive students (20.8 ± 3.7 years; 50% female; 97% nonsmokers) from the university. They were categorized into four conditions: natural and task group, urban and task group, natural and no-task group, and urban and no-task group</td>
</tr>
<tr>
<td>Intervention</td>
<td>20-minute walking exercise for each condition. Participants were advised to exercise at level 12 (fairly light) of Borg’s 20-point rating of perceived exertion scale. The speed was controlled remotely by the tester via a software package according to oral feedback from participants. Four intervention groups of 20 participants were exposed to a sequence of 30 scenes projected on a wall whilst exercising on a treadmill. Four categories of scenes were tested: rural pleasant, rural unpleasant, urban pleasant, and urban unpleasant</td>
<td>To vary restoration needs, half of the participants began the environmental treatment directly after driving to the field site (no-task group). The other half completed attention demanding tasks just before the treatment (task group). After the drive or the task, half of each (no-task, task) group sat in a room with a view of a tree and then walked in the natural environment. The other half of each group sat in a room without a tree view and then walked in the urban environment</td>
</tr>
<tr>
<td>Main and secondary outcomes</td>
<td>Blood pressure and two psychological measures (self-esteem and mood)</td>
<td>Blood pressure, emotion, and attention</td>
</tr>
<tr>
<td>Main results</td>
<td>Exercise alone significantly reduced blood pressure, increased self-esteem, and had a significant positive effect on mood measures. Both rural and urban pleasant scenes produced a significantly greater positive effect on self-esteem than the exercise-only control</td>
<td>After the drive or the task, sitting in a room with a tree view promoted more rapid decline in diastolic blood pressure than a viewless room. Subsequently walking in the nature reserve initially fostered blood pressure change that indicated greater stress reduction than afforded by walking in the urban surroundings. Performance on an attentional test improved slightly from the pretest to the midpoint of the walk in the nature reserve, while it declined in the urban setting. Positive effect increased and anger decreased in the nature reserve by the end of the walk; the opposite pattern emerged in the urban environment</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Green exercise has important public and environmental health consequences</td>
<td>Public health strategies with a natural environment component may have a particular value in this time of growing urban populations, exploding health care expenditures, and deteriorating environmental quality</td>
</tr>
</tbody>
</table>

(Continued)
“for each primary and secondary outcome, results for each group, and the estimated effect size and its precision;” “all important harmful or unintended effects in each group;” “trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses;” “registration number and name of trial registry;” and “where the full trial protocol can be accessed, if available.”

Table 3 presents an assessment of the evaluation of study quality using the CLEAR NPT checklist. The CLEAR NPT checklist is a tool mainly used to evaluate the quality of the study conduct. The items in which the description was lacking in both studies were as follows: “was the treatment allocation concealed?;” “were participants adequately blinded?;” and “were outcome assessors adequately blinded to assess the primary outcomes?”,

**Discussion**

Only two RCTs on forest therapy were identified, which indicates that there is little evidence demonstrating the effectiveness of the therapy on mental health improvement. One reason for the limited number of RCTs might be that there are few experimenters in forest therapy who are familiar with the methodology or study design of interventions. Another possibility is the innate difficulty of study blinding; participants inevitably notice the intervention, which makes RCT design difficult.

Furthermore, in this study it was impossible to perform a meta-analysis and integrate the results as the main outcome measures and interventions were different and could not be compared between the papers. This suggests that more high-quality studies like RCTs on forest therapy are needed in order to build the evidence, no matter how difficult it is.

**Overall evidence and quality assessment**

The CONSORT 2010 and CLEAR NPT checklists were used as quality assessments. There were serious problems with the conduct and reporting of the target articles. The summaries detected omissions of description, including identification as a randomized trial in the title, method used to generate the random allocation sequence, blinding, estimated effect size and its precision, trial limitations, addressing sources of potential bias, and analysis methods.

In the Cochrane Reviews, the eligibility criteria for a meta-analysis are strict, and for each article, heterogeneity and low quality of reporting must first be excluded. Because there was insufficient evidence in the studies of forest therapy, due to poor methodological and reporting quality and heterogeneity, it was not possible to offer any conclusions about the effects of forest therapy based on a systematic review.

**Characteristics of articles and other studies**

**Short duration experiment**

The two studies that met all inclusion criteria were of short duration experiments (20–60 minutes). Effects over the long term were not able to be determined. Many earlier studies (non-RCTs) are short-term experiments, and observe the change on trips of only a few days, if that long.6,7

**Setting and interventional environment**

Pretty’s study was the experiment by virtual reality, and Hartig et al.’s study were performed under actual environment. There are many studies in which urban, city, or man-made environments are used as a contrast for the intervention.4,11–16 We assumed that the setting of the control is the key to demonstrating the superiority of the forest therapy.

**Setting**

There were very few RCT intervention studies on forest therapy, and most of the candidate studies that were detected were crossover design studies (Figure 1). Crossover trials are easier to carry out with small sample sizes than parallel group trials. However, researchers should suspect that there is a strong preconception that nature, green surroundings, and forest environments are good for the mind and body, and that it is very likely that results may be biased in the direction that the author aims the study.

**Future research agenda**

Table 4 shows the future research agenda for forest therapy. Researchers should use the appropriate checklists for research design and intervention method, which would lead to an improvement in the quality of the study, and contribute to demonstrating the superiority of the forest therapy.
Table 2 Evaluation of the quality of randomized controlled trials by using the Consolidated Standards of Reporting Trials 2010 checklist

<table>
<thead>
<tr>
<th>Paper section/topic</th>
<th>ID</th>
<th>Checklist item</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td>1a</td>
<td>Identification as a randomized trial in the title</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>1b</td>
<td>Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)</td>
<td>No</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>2a</td>
<td>Scientific background and explanation of rationale</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>Specific objectives or hypotheses</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>3a</td>
<td>Description of trial design (eg, parallel, factorial) including allocation ratio</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>Important changes to methods after trial commencement (eg, eligibility criteria), with reasons</td>
<td>No</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>4a</td>
<td>Eligibility criteria for participants</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>Settings and locations where the data were collected</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Interventions</strong></td>
<td>5</td>
<td>The interventions for each group with sufficient details to allow replication, including how and when they were actually administered</td>
<td>No</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>6a</td>
<td>Completely defined prespecified primary and secondary outcome measures, including how and when they were assessed</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>6b</td>
<td>Any changes to trial outcomes after the trial commenced, with reasons</td>
<td>NR</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>7a</td>
<td>How sample size was determined</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>7b</td>
<td>When applicable, explanation of any interim analyses and stopping guidelines</td>
<td>No</td>
</tr>
<tr>
<td><strong>Randomization:</strong></td>
<td>8a</td>
<td>Method used to generate the random allocation sequence</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>8b</td>
<td>Type of randomization; details of any restriction (such as blocking and block size)</td>
<td>No</td>
</tr>
<tr>
<td><strong>Allocation concealment mechanism</strong></td>
<td>9</td>
<td>Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned</td>
<td>No</td>
</tr>
<tr>
<td><strong>Implementation</strong></td>
<td>10</td>
<td>Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions</td>
<td>No</td>
</tr>
<tr>
<td><strong>Blinding</strong></td>
<td>11a</td>
<td>If done, who was blinded after assignment to interventions (eg, participants, health care providers, those assessing outcomes) and how</td>
<td>NB</td>
</tr>
<tr>
<td></td>
<td>11b</td>
<td>If relevant, description of the similarity of interventions</td>
<td>NR</td>
</tr>
<tr>
<td><strong>Statistical methods</strong></td>
<td>12a</td>
<td>Statistical methods used to compare groups for primary and secondary outcomes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>12b</td>
<td>Methods for additional analyses, such as subgroup analyses and adjusted analyses</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>13a</td>
<td>For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>13b</td>
<td>For each group, losses and exclusions after randomization, together with reasons</td>
<td>–</td>
</tr>
<tr>
<td><strong>Recruitment</strong></td>
<td>14a</td>
<td>Dates defining the periods of recruitment and follow-up</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>14b</td>
<td>Why the trial ended or was stopped</td>
<td>–</td>
</tr>
<tr>
<td><strong>Baseline data</strong></td>
<td>15</td>
<td>A table showing baseline demographic and clinical characteristics for each group</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Numbers analyzed</strong></td>
<td>16</td>
<td>For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Outcomes and estimation</strong></td>
<td>17a</td>
<td>For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (eg, 95% confidence interval)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>17b</td>
<td>For binary outcomes, presentation of both absolute and relative effect sizes is recommended</td>
<td>NR</td>
</tr>
<tr>
<td><strong>Ancillary analyses</strong></td>
<td>18</td>
<td>Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing prespecified from exploratory</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Harms</strong></td>
<td>19</td>
<td>All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)</td>
<td>No</td>
</tr>
<tr>
<td><strong>Discussion</strong></td>
<td>20</td>
<td>Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses</td>
<td>No</td>
</tr>
<tr>
<td><strong>Generalizability</strong></td>
<td>21</td>
<td>Generalizability (external validity, applicability) of the trial findings</td>
<td>No</td>
</tr>
<tr>
<td><strong>Interpretation</strong></td>
<td>22</td>
<td>Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence</td>
<td>No</td>
</tr>
</tbody>
</table>

(Continued)
Table 2 (Continued)

<table>
<thead>
<tr>
<th>Paper section/topic</th>
<th>ID</th>
<th>Checklist item</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other information</td>
<td></td>
<td></td>
<td>11 12</td>
</tr>
<tr>
<td>Registration</td>
<td>23</td>
<td>Registration number and name of trial registry</td>
<td>No  No</td>
</tr>
<tr>
<td>Protocol</td>
<td>24</td>
<td>Where the full trial protocol can be accessed, if available</td>
<td>No  No</td>
</tr>
<tr>
<td>Funding</td>
<td>25</td>
<td>Sources of funding and other support (eg, supply of drugs), role of funders</td>
<td>No  Yes</td>
</tr>
</tbody>
</table>

Abbreviations: CONSORT, Consolidated Standards of Reporting Trials; NB, no blinding; NR, not relevant.

Table 3 Evaluation of the quality of randomized controlled trials by using the CLEAR NPT (A checklist to evaluate a report of a nonpharmacological trial) checklist

<table>
<thead>
<tr>
<th>Items</th>
<th>Practice</th>
</tr>
</thead>
</table>
| 1. Was the generation of allocation sequences adequate?  
Reference 11  
Reference 12                                                                                                                                  | Yes  o    |
| 2. Was the treatment allocation concealed?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 3. Were details of the intervention administered to each group made available?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 4. Were care providers’ experience or skill in each arm appropriate?  
Reference 11  
Reference 12                                                                                                                                  | Not relevant  Not relevant |
| 5. Was participant (ie, patients) adherence assessed quantitatively?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 6. Were participants adequately blinded?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 6.1. If participants were not adequately blinded,  
6.1.1. Were all other treatments and care (ie, cointerventions) the same in each randomized group?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 6.1.2. Were withdrawals and lost to follow-up the same in each randomized group?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 7. Were care providers or persons caring for the participants adequately blinded?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 7.1. If care providers were not adequately blinded,  
7.1.1. Were all other treatments and care (ie, cointerventions) the same in each randomized group?  
Reference 11  
Reference 12                                                                                                                                  | Not relevant  Not relevant |
| 7.1.2. Were withdrawals and lost to follow-up the same in each randomized group?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 8. Were outcome assessors adequately blinded to assess the primary outcomes?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 8.1. If outcome assessors were not adequately blinded, were specific methods used to avoid ascertainment bias (systematic differences in outcome assessment)?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |
| 9. Was the follow-up schedule the same in each group?  
Reference 11  
Reference 12                                                                                                                                  | Not relevant  Not relevant |
| 10. Were the main outcomes analyzed according to the intention-to-treat principle?  
Reference 11  
Reference 12                                                                                                                                  | o  o      |

Notes:  
*First and second items were not described in order for the randomized controlled trial design; ‘the answer should be “yes” for this item if these data were either described in the report or made available for each arm (eg, reference to a preliminary report, online addendum); ‘care provider experience or skill will be assessed only for therapist-dependent interventions (ie, interventions where the success of the treatment are directly linked to care provider’s technical skill). For other treatment, this item is not relevant and should be removed from the checklist or answered “unclear;” ‘appropriate experience or skill should be determined according to published data, preliminary studies, guidelines, run-in period, or a group of experts and should be specified in the protocol for each study arm before the beginning of the survey; ‘treatment adherence will be assessed only for treatments necessitating interventions (eg, physiotherapy that supposes several sessions in contrast to a one-shot treatment such as surgery). For one-shot treatments, this item is not relevant and should be removed from the checklist or answered “unclear;” ‘the answer should be “yes” for this item, if the main outcome is objective or haed, or if outcomes were assessed by a blinded, or at least an independent, endpoint review committee, or if outcomes were assessed by an independent outcome assessor trained to perform the measurements in a standardized manner, or if the outcome assessor was blinded to the study purpose and hypothesis; ‘this item is not relevant for trials in which follow-up is part of the question. For example, this item is not relevant for a trial assessing frequent versus less frequent follow-up for cancer recurrence. In these situations, this item should be removed from the checklist or answered “unclear.”
to the accumulation of evidence. Appropriate comparisons are necessary to explain why forest therapy is better than other types of interventions. Long-term intervention studies are necessary to clarify the effect of forest therapy. Researchers should not only present the efficacy data, but also any adverse events or harmful phenomena. For example, people with pollen allergies refuse this treatment in some seasons. In addition, it may be necessary to add original items such as herbal intervention, aquatic exercise, and balneotherapy to the CONSORT checklist as alternative medicines.

A recent study suggested that public health is moving toward the goal of implementing evidence-based intervention. But the feasibility of possible interventions and whether comprehensive and multilevel evaluations are needed to justify them must be determined. It is at least necessary to show cost-benefit relations.

In addition, a well-designed observational study to clarify that people living in the forest or rural areas are healthy is indispensable. Furthermore, researchers must definitely answer the question, “Are forestry workers really healthy?” A recent cohort study reported that populations exposed to greener environments also enjoy lower levels of income-related health inequality, and, conversely, populations exposed to less green environments could be less protected from health inequality related to income deprivation, which might have ramifications for countries in which urbanization remains a strong force.

**Table 4 Overall evidence and future research agenda to build evidence on forest therapy**

<table>
<thead>
<tr>
<th>Overall evidence in the present</th>
<th>Research agenda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor/unclear (forest specific effect is not clear)</td>
<td><em>Randomized controlled trial:</em> Structural description of papers based on CONSORT 2010</td>
</tr>
<tr>
<td></td>
<td>1. Satisfactory description and methodology (identification as a randomized trial in the title, method used to generate the random allocation sequence, blinding, estimated effect size and its precision, trial limitations, addressing sources of potential bias, and analysis methods)</td>
</tr>
<tr>
<td></td>
<td>2. Setting of the appropriate control area or place</td>
</tr>
<tr>
<td></td>
<td>3. The long-term effects of forest environment</td>
</tr>
<tr>
<td></td>
<td>4. Description of adverse effects (eg, pollen allergy)</td>
</tr>
<tr>
<td></td>
<td>5. The development of the original check item in forest therapy</td>
</tr>
<tr>
<td></td>
<td>6. Cost-benefit (eg, travel expense and medical cost)</td>
</tr>
</tbody>
</table>

**Study limitations**

This study was based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement, except for the meta-analysis. However, there were several limitations to the study. Some selection criteria were common across studies, as described above; however, bias remained due to differences in eligibility for participation in each study. Publication bias was also a limitation. Although there was no linguistic restriction in the eligibility criteria, only studies with English and Japanese keywords were searched. Furthermore, the references could not be checked by handsearching, and proof of the mechanisms in the clinical trials was not evaluated. The Allied and Complementary Medicine Database, a bibliographic database (several professions allied to medicine, complementary medicine, and palliative care covered a selection of journals in three separate subject areas) produced by the Health Care Information Service of the British Library, was not searched.

**Conclusion**

Because there was insufficient evidence on forest therapy due to poor methodological and reporting quality and heterogeneity of RCTs, it was not possible to offer any conclusions about the effects of this type of intervention. However, it was possible to identify the problems with current RCTs of forest therapy, and to propose a strategy for strengthening study quality and stressing the importance of study feasibility and original checklist items based on characteristics of forest therapy as a future research agenda objective.

**Acknowledgments**

The authors would like to express their appreciation to Ms Rie Higashino (paperwork) and Ms Mari Makishi (all database/registry searches) for their assistance in this study. This study was supported by the Nakayama Foundation for Human Science Research Grant 2010.

**Disclosure**

The authors report no conflicts of interest in this work. HK and SJP conceived the study and take responsibility for the quality assessment and summary of included studies and data extraction. KT and YM are the guarantors. HK, SO, HO, and SH designed the study. NS, JK, TH, and HK acquired the data. MK, SJP, TH, and HK assessed the quality of articles. All authors critically revised the manuscript for important intellectual content.
References


### Appendix

References to studies excluded in this review

<table>
<thead>
<tr>
<th>Exclusion number</th>
<th>Citation</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Takayanagi K, Hagiwara Y. To extend health resources in a forest hospital environment: a comparison between artificial and natural plants. J Jpn Mibyo Syst Assoc. 2006;11(2):247–259.</td>
<td>Nonrandomized controlled trial</td>
</tr>
</tbody>
</table>