Systematic review of the effectiveness of aromatherapy in labor

Hitomi Suzuki  
St. Luke's International University  
https://orcid.org/0000-0002-1747-7601

Mika Kato  
Sei Roka Kokusai Daigaku

Katsuko Shinozaki  
Kokusai Iryo Fukushi Daigaku

Katharina da Silva Lopes  
Sei Roka Kokusai Daigaku

Erika Ota  
ota@slcn.ac.jp  
https://orcid.org/0000-0002-3945-7441

Research article

Keywords: Aromatherapy, Labor pain, Anxiety, Complementary therapies

Posted Date: August 4th, 2019

DOI: https://doi.org/10.21203/rs.2.12405/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. 
Read Full License
Abstract

Background Clinical practice with aromatherapy has become an expanding area for nursing, and is considered one of the most popularly used complementary treatments. However, there is insufficient evidence about the benefits of aromatherapy for pain management and other related discomforts in labor. We aimed to evaluate the effects of aromatherapy for women during delivery particularly for pain relief. Methods AMED, ClinicalTrials.Gov, CINAHL, Cochrane Library, EMBASE, MEDLINE, PubMed, and WHO ICTRP were searched in August 2017. For updates, these databases were searched from July 2017 to July 2018. This study included randomized controlled trials (RCTs) and quasi-RCTs for normal pregnancy women who were experiencing labor onset, and compared aromatherapy with standard care or control. Results Six RCTs from six reports, and four quasi-RCTs from five reports were included (1238 pregnant women). The trials found significant difference between groups for the primary outcomes of pain relief on the latent (MD -1.56, 95%CI: -2.45 to -0.67, low certainty of evidence) and early active phase (MD -1.69, 95%CI: -2.50 to -0.89, low certainty of evidence). However, there were no significant differences for the primary outcomes of pain relief on the late active phase, and anxiety relief on the early and late active phases. Conclusions This meta-analysis found evidence that the use of inhalation aromatherapy for term pregnancy women is associated with reduction of labor pain. However, there is insufficient evidence to confirm pain relief on the late active phase, anxiety relief, and other outcomes following aromatherapy. Trial registration We registered the study protocol with PROSPERO (CRD42017077617).

Background

Although labor and birth are considered to be a natural process, laboring women experience a significant amount of discomfort and pain as well as a variety of other challenging sensations [1]. Women in labor experience pain caused by uterine contractions, expansion of the lower uterus, and the dilation of the cervix. Moreover, pain is produced by the stretching of the vagina and pelvic floor to accommodate the baby. These complexes of pain can lead to complications such as a compromised immune system, reactive hypoglycemia, delayed wound healing, increased myocardial oxygen consumption, paralytic ileus, and reduced respiratory function [2]. This pain possibly exerts its effects in the form of psychological distress to both the mother and the baby. Perceptions of labor pain vary by the individual and are influenced by a variety of physiologic, psychologic, emotional, socio-cultural, and environmental factors [1]. Labor, without using drugs or invasive methods such as an epidural, is often sought by many women and they usually turn to complementary therapies such as aromatherapy to help reduce their pain perception [3].

Aromatherapy is the use of essential oils from plants such as flowers, trees or herbs. These essential oils improve physical, mental, and spiritual well-being. The clinical practice with aromatherapy has become an expanding area for nursing and is considered one of the most popularly used complementary therapies [4]. Aromatherapy during labor and delivery may provide relaxation and reduce pain [5]. Several essential oils used in aromatherapy have been suggested to have antistressor, antidepressive, and
relaxation effects [6]. In addition to reducing pain, aromatherapy may also decrease symptoms such as anxiety, nausea, vomiting and other labor-related conditions [6].

To date, there has been inconclusive evidence regarding the benefits of aromatherapy for the management of pain and other related discomforts in labor, as well as for the improvement of maternal and neonatal outcomes. A previous Cochrane systematic review on aromatherapy published in 2011 could not clearly show evidence of its effects on pain relief during deliveries [3]. The most recent systematic review, which was published in 2019, reported the anxiolytic effects of aromatherapy, and suggested positive effects on anxiety during the first stage of labor [7]. This 2019 review analyzed the efficacy of individual aroma essence oils in the first stage of labor, and it included studies of aromatherapy intervention using other care methods such as massage. We concluded that aromatherapy results may be influenced by other factors.

We believe that a meta-analysis is needed to definitively evaluate the efficacy of inhalation aromatherapy in terms of all its outcomes, particularly for pain management by excluding cointervention, which may affect its efficacy. The purpose of this systematic review is to evaluate the efficacy of aromatherapy for women in all stages of labor.

**Methods**


**PICOS and selection criteria**

We included all types of randomized controlled trials (RCTs) involving pregnant women of 37-42 gestation weeks and with labor onset. The included RCTs only focused on inhalation of any kind of aroma essence compared with the standard care or placebo. We included multiple arms and cointervention, which were to compare the efficacy of aromatherapy. The primary outcomes were pain relief and anxiety relief during labor. The secondary outcomes were duration of delivery, duration of contraction, spontaneous or operative delivery, and Apgar score (Appendix S1).

**Search strategy**

In August 2017, we searched AMED, ClinicalTrials.Gov, CINAHL, Cochrane Library, EMBASE, MEDLINE, PubMed, and WHO ICTRP with no date/time, language, document type, and publication status limitations. For the present review, these databases were searched from July 2017 to July 2018. The PubMed search strategy is shown in the Appendix S2, and we adapted the search terms for every database. We also hand searched Google scholar in August 2018.

**Data extraction**
KS, MK, and HS independently screened and confirmed study eligibility. When there were conflicts of eligibility, each study was discussed with HS and EO for resolution. We found several papers written in Persian, and we attempted to contact to the authors. However, we received no reply. From each study, KS extracted information on characteristics of participants, study design, numbers of participants, interventions, and outcome data. Data were extracted by HS and checked by MK. Risk of bias was assessed as recommended in the Cochrane Handbook [8], and MK and HS independently assessed each trial. Discrepancies were resolved through discussion. MK and HS also contacted the authors of three RCTs [12-14] to request unpublished outcome data, where trial reports implied that relevant data might be available. However, replies were not forthcoming.

**Statistical analysis**

We performed meta-analysis to analyze pooled outcome data. We estimated weighted mean difference (MD) when the outcomes were measured similarly between trials. We also used standardized mean difference (SMD) to combine trials that measured the same outcome with different methods, and 95% confidence interval (CI) for continuous outcomes. If multiple intervention arms were reported, we combined the data of aromatherapy intervention groups and calculated mean and standard deviation (SD). When cointervention was used, we analyzed the compared data of aromatherapy and controlled to avoid the effects of other interventions. For binary outcomes (e.g., response, remission, and dropouts), we estimated risk ratio (RR) and 95% CI for each comparison using the numbers randomly assigned and numbers of events. We used intention-to-treat (ITT) data in this analysis, as ITT data are less biased and address a more pragmatic and clinically relevant situation. To address missing data, we used the number randomized minus any participants for the denominator for each outcome in each trial. We used general inverse variance method (GIVM) when the included studies reported only the difference between the means for the two groups and the standard error of this difference. We included outcome data from quasi-RCTs. We carried out sensitivity analysis to explore the effect of trial quality assessed by quasi-RCTs, concealment of allocation and incomplete outcome data, or more than one, with quasi-RCT studies being excluded from the analysis to assess whether this makes any difference in the overall result. Heterogeneity was assessed using $I^2$ statistic. We considered $I^2 \geq 60\%$ as high, then we used random effects meta-analysis. For low to moderate heterogeneity ($I^2 < 60\%$), we used fixed effects meta-analysis.

We performed all analyses using Review Manager (RevMan) [15]. We used GRADE [16] to judge the certainty of evidence for the effectiveness of aromatherapy for the primary outcomes such as labor pain relief and anxiety relief through all the stages of labor.

**Results**

**Trial characteristics**

We screened 254 titles and abstracts (Fig. 1), and identified six individual RCTs from six reports [18, 19, 22, 25-27] and four quasi-RCTs from five reports [17, 20, 21, 23, 24] for inclusion in the final review. We
excluded the trials with massage and bathing as these have other effects. Eight studies that were included mentioned that blinding of participants was not possible because of the diffusion of oil molecules in the air [17, 19-22, 25-27]. For this reason, five reports [17, 20, 21, 23, 24] had performed interventions on randomly allocated days with the aromatherapy days and the placebo days. Although we considered these four trials as quasi-RCTs, we included these four trials from the nature of the intervention.

Table 1 provides details of the included studies involving 1238 pregnant women at labor onset [17-27]. Eight trials included only nulliparous women (81.0%) [17,20-27], one trial did not report parity (8.9%) [18], and one trial showed the mean ± SD of the numbers of parity 1.31 ± 0.72 for the intervention group and 1.22 ± 0.91 for the control group (9.6%) [19]. All of the trials recruited participants with singleton pregnancy and full-term pregnancy, and did not report existing medical conditions. Most trials recruited predominantly adults (18-35 years old) with cephalic presentation and 3-4 cm cervical dilatation. Nine trials were undertaken in Iran [17-25, 27] and only one was conducted in Thailand [26]. All of the trial settings were at hospitals.

For the measurement of outcomes, labor pain severity was measured using the Visual Analogue Scale (VAS) chart and the Numerical Rating Scale (NRS) [17-21, 24, 26, 27]. Both scales have a score range of 0 to 10 [28, 29]. One trial reported the pain score changes from baseline therefore we performed GIVM for the meta-analysis [26]. In three studies, Spielberger’s State-Trait Anxiety Inventory (STAI) was used to determine the level of anxiety of the participants [18, 23, 25]. STAI questionnaires consist of 40 questions in which the scores ranged from 20 to 80. Higher scores indicate greater anxiety [30]. The reliability of STAI has a Cronbach’s alpha of 0.90 [25]. One study used the Visual Analog Scale for Anxiety (VASA) [22]. The scale ranges from 0 to 10 with 0 indicating no anxiety and 10 greatest anxiety [31].

Risk of bias for included studies

Of the 10 trials, most of the trials had a low risk of bias in random sequence generation (60%, 6/10), incomplete outcome data (90%, 9/10), selective reporting (60%, 6/10), and other bias (100%, 10/10). However, most of the trials had a high risk or unclear bias in allocation concealment (60%, 6/10), blinding of participants and personnel (100%, 10/10), and blinding of outcome assessment (60%, 6/10). Eight studies mentioned that blinding participants were not possible owing to the nature of aromatherapy [17, 19-22, 25-27]. Four trials from five reports [17, 20, 21, 23, 24] had performed interventions on randomly allocated days; thus, we considered these four trials as quasi-RCTs.

Aromatherapy interventions

Table 1 presents details of the aromatherapy interventions administered in each trial. All trials evaluated inhalation of aroma essence in labor. Two studies had a three-arm design with intervention arms [21, 22]. One study used two kinds of aroma essence (Jasmin and Salvia essence), and we combined them into one group and used the calculated data which is the combined mean ±SD [21]. Another study carried out interventions by inhalation of aroma essence using a footbath, only footbath, and routine care [22]. We
included inhalation of aroma essence with footbath as the intervention group, and only footbath as the control group to exclude the effect from footbath. Moreover, one study performed inhalation of aroma essence with breath technique as the intervention, and breath technique alone as the control [19].

Various aroma essences were used in the included trials. Lavender was the most used aroma oil in four trials [19, 20, 26, 27], and it is also commonly used in practice settings. The second most used aroma essences were C. aurantium essence [23, 24, 26], Geranium rose essence [25, 26], Jasmin [21, 26], and R. damascene essence [18, 22] in two trials each, and Salvia essence [21] and Boswellia carterii essence [17] were used in single trials each.

**Primary outcomes**

*Labor pain relief*

For the measurement of labor pain, all of the studies used VAS or NRS [17-21, 24, 26, 27]. As one trial reported interquartile range we used score change reports [26]. For this reason, we calculated MD with GIVM for the analysis of labor pain. Eight studies found that aromatherapy significantly reduced labor pain intensity compared with control in the latent phase (MD -1.56, 95% CI -2.45 to -0.67, p = 0.0006, I² = 97%, eight trials, 1,005 women, low certainty of evidence; Fig 3). Six studies reported that aromatherapy intervention significantly reduced labor pain compared with control in the early active phase (MD -1.69, 95% CI -2.50 to -0.89, p < 0.0001, I² = 96%, six trials, 689 women, low certainty of evidence; Fig. 4). These studies also reported that aromatherapy significantly reduced labor pain in the late active phase (MD -1.52, 95% CI -2.33 to -0.71, p = 0.0002, I² = 97%, six trials, 689 women, very low certainty of evidence; Fig. 5).

*Anxiety relief*

Three studies used STAI [18, 23, 25] and one study used VASA [22] to measure the outcome of anxiety. Therefore, we calculated SMD for the analysis of anxiety. Aromatherapy intervention reduced anxiety compared with the control in the early active phase (SMD -3.49, 95% CI -6.28 to -0.69, p = 0.01, I² = 99%, four studies, 392 women, very low certainty of evidence; Fig. 6). Three studies reported that aromatherapy significantly reduced anxiety in the late active phase. (SMD -5.54, 95% CI -10.39 to -0.69, p = 0.03, I² = 99%, three studies, 295 women, very low certainty of evidence; Fig. 7).

**Secondary outcomes**

*Duration of contraction*

We used SMD for the analysis of duration of contraction because the time unit of the included studies was unclear [17, 18, 23]. Three studies found that aromatherapy did not significantly affect the duration of contractions at 3-4 cm, 5-7 cm, and 8-10 cm dilatation (3-4 cm; SMD -0.49, 95% CI -1.41 to 0.43, p = 0.30, I² = 94%, 347 women; Appendix. Fig. S1), (5-7 cm; SMD 2.94, 95% CI -0.38 to 6.26, p = 0.08, I² = 99%,
347 women; Appendix. Fig. S2), (8-10 cm; SMD 0.05, 95% CI -0.16 to 0.26, p = 0.67, \( l^2 = 49\% \), 347 women; Appendix. Fig. S3).

**Labor length**

We also used SMD for the analysis of labor length as the time unit was unclear [19-21, 26, 27]. Five studies showed that aromatherapy significantly reduced the 1\(^{st}\) stage labor length compared with the control (SMD -0.21, 95% CI -0.37 to -0.06, \( p = 0.008 \), \( l^2 = 56\% \), five studies, 641 women; Appendix. Fig. S4). By contrast, aromatherapy did not significantly reduce the 2\(^{nd}\) stage labor length (SMD 0.14, 95% CI -0.36 to 0.63, \( p = 0.59 \), \( l^2 = 86\% \), four studies, 481 women; Appendix. Fig. S5).

**Apgar score**

Aromatherapy intervention did not significantly affect the Apgar score at 1 min after delivery (MD -0.25, 95% CI -0.84 to 0.35, \( p = 0.41 \), \( l^2 = 98\% \), five studies, 652 women; Appendix. Fig. S6) (17,18,20,21,27). These five studies also report that aromatherapy was not significantly associated with the Apgar score at 5 min (MD 0.03, 95% CI -0.02 to 0.08, \( p = 0.25 \), \( l^2 = 0\% \), five studies, 652 women; Appendix. Fig. S7).

We calculated RR for dichotomous results. Only one study reported Apgar score < 7 at 1 min [26]; however, there was no significant association with aromatherapy (RR 0.51, 95% CI 0.05 to 5.45, \( p = 0.58 \), one study, 103 women; Appendix. Fig. S8). An Apgar score < 7 at 5 min was not reported.

**Types of delivery**

Three studies reported that aromatherapy intervention did not increase spontaneous delivery [18, 21, 26] (RR 1.05, 95% CI 0.96 to 1.15, \( p = 0.27 \), \( l^2 = 0\% \), three studies, 370 women; Appendix. Fig. S9). Aromatherapy did not significantly reduce operative delivery [18, 26] (RR 0.58, 95% CI 0.24 to 1.42, \( p = 0.24 \), \( l^2 = 0\% \), two studies, 214 women; Appendix. Fig. S10), or C-sections [18,21,23,26] (RR 0.83, 95% CI 0.47 to 1.49, \( p = 0.54 \), \( l^2 = 0\% \), four studies, 483 women; Appendix. Fig. S11).

**Labor augmentation**

Only one study reported the ratio of labor augmentation [26]. Aromatherapy did not significantly reduce labor augmentation (RR 0.97, 95% CI 0.68 to 1.37, \( p = 0.84 \), one study, 104 women; Appendix. Fig. S12).

**Sensitivity analysis**

Due to the high heterogeneity, we performed sensitivity analysis by excluding quasi-RCTs and high risk of random sequence [17, 20, 21, 23, 24]. For labor pain relief, it still showed significant differences during the latent and early active phase. (Appendix. Fig. S13, S14). However, there were no significant differences in pain relief during the late active phase, and anxiety relief during all of the active stages (Appendix. Fig.S15-S17). For secondary outcomes, there was no significant difference in the labor length of the 1\(^{st}\) stage (Appendix. Fig. S18).
Discussion

We found evidence that using inhalation aromatherapy for term pregnancy women was associated with reduction of labor pain on the latent and early active phase. This is most likely the first review showing evidence of pain relief with the use of aromatherapy in labor, and the first study to evaluate the efficacy of aromatherapy in all the stages of labor. The Cochrane systematic review included two RCTs [32, 33]. However, the researchers could not perform a meta-synthesis owing to the differences in the comparison methods, and they reported each trial result individually. This previous review could not identify any evidence of the effects of aromatherapy on pain relief in labor because the number of included studies were not sufficient. Furthermore, the searches were conducted in 2010 [3]. In the present review, there was insufficient evidence regarding the effect of reducing anxiety of the participants using inhalation aromatherapy during labor. Ghiasi et al. reported the systematic review on the anxiolytic effect of aromatherapy during the first stage of labor [7]. They suggested a positive effect on anxiety with a qualitative analysis of the benefit from individual aroma essence oils in the first stage of labor. This previous review also did not perform a meta-analysis, thus this review was not conclusive [7]. Moreover, Ghiasi et al. included trials of aromatherapy intervention with massage. In the present review, we excluded data that may exert any effect on the evaluation of the genuine efficacy of aromatherapy. Thus, our hypothesis that the use of aromatherapy during labor reduces labor pain was found to be effective in pregnant women on the latent and early active phase, and we considered a low certainty of evidence from the GRADE assessment (Table 2). The Cochrane review reported outcomes of assisted vaginal delivery, cesarean delivery, spontaneous delivery, augmentation, and admission to NICU, and there was no evidence of effect due to the lack of power [3]. Taken together, there were no significant differences in the secondary outcomes, such as duration of contraction, labor length, Apgar score, types of delivery, labor augmentation in our present review. Moreover, there was no report about adverse events of aromatherapy in the included studies. However, people have different preferences of smell, and pregnant women are particularly more sensitive. The choice of essential oils depends on the women's preference.

Overall, the risk of bias about blinding allocation of the included studies were high. Blinding allocations of the participants and care providers are difficult because of the smell of the aroma essence. This setting downgrades the level of evidence. Four trials of random sequence also had a high risk of bias. We conducted sensitivity analysis to adjust for the effects of this high risk of bias of randomization. The results showed high heterogeneities overall in the outcomes of pain as well as reduction of anxiety. In another review, labor pain intensity was also reported with high heterogeneity [34]. Pain and anxiety are subjective senses, and these outcomes were assessed by self-report questionnaires. Additionally, it is predictable that there are various individual differences in sensitivity to labor pain. These various differences may provoke a high heterogeneity status.

We did not specify the kinds of aroma essence although eight kinds of essences were used in the included studies. We cannot define the specific efficacy of aromatherapy in this present review. It is
possible that what influenced the reduction in the subjective labor pain was the efficacy of relaxation brought about by the inhalation of pleasant smells of the aroma essence. A woman's internal experience of labor pain is affected by the environment, and this factor includes the person's verbal and nonverbal communications [35]. Relaxation may have a role in reducing pain, increasing satisfaction with pain relief, and reducing the rate of assisted vaginal delivery [36]. Creating an environment with less stress and providing relaxation to parturient women are needed to reduce pain intensity.

Moreover, there is a strong association between anxiety and pain in the latent phase of labor [37]. Although we could not find sufficient evidence of reducing anxiety by aromatherapy, the results showed a tendency for aromatherapy to be slightly more effective than the control group. From this point of view, even if we could not reach definitive evidence of anxiety relief in this present review, there were still benefits of using aromatherapy in labor.

For women, natural pain and anxiety interventions are in more demand than medical interventions. Aromatherapy is a noninvasive method with a low risk and low cost for reducing labor pain.

**Strength and Limitations**

This present review has several strengths. Almost all included studies used genuine branded aroma oils for women in labor, thus it is sufficient to say that most of the outcomes are reliable in terms of their efficiency. Only one study had an unclear attrition bias [22]. This included study reported the number of assigned participants, but not the number of randomized participants. However, other included reports all had a low risk of bias in *incomplete outcome data*, and also *other bias* was low in this review. From this status, we considered that the quality of evidence in the present review was kept from attrition bias. This systematic review provides the results of all outcomes following aromatherapy in all stages of labor. Although we could not assess the differences in the efficacies of individual essential oils, Ghiasi et al. reported the efficacy of aromatherapy on anxiety for each aroma oil [7]. Investigating differences in the effectiveness of essential oils on labor pain should also be considered.

Despite these strengths, several limitations were found in the present review. Firstly, the nine of the 10 studies for inclusion in the final review were conducted in Iran. The results of these studies might be impacted by its bias. Secondly, all studies could not maintain blinding of participants because the essential oil fragrance naturally spreads in the air and prevents from complete blinding of the participants. Additionally, this study was able to examine the effectiveness of aromatherapy, and found a possible effect for reducing women's labor pain. Although there was a positive result for pain relief, it is very difficult to evaluate this objectively because the perception of pain is very subjective and results from various internal experiences. Women in labor experience a significant amount of discomfort rather than labor pain, but the included studies were not focused on those unpleasant symptoms such as nausea, excessive physical sensitivity, vomiting, and other discomforts.

Future research is needed not only in terms of physiological labor pain but also psychological efficacy (e.g., stress response changes or anxiety). The former Cochrane review concluded that the efficacy and
effectiveness of aromatherapy have not yet been established owing to the limited number of trials [3]. Our included studies were comparatively new, and we updated the efficacy for pain reduction.

Further research of the evaluation of aromatherapy about these three points are needed. Firstly, the included studies were concentrated in Iranian settings, and future research should be investigated in other various settings. Secondly, the types of aroma essence were not specified in this review, and future research should assess the efficacy of the specific kinds of aroma oils. Thirdly, future research should investigate additional outcomes such as nausea, excessive physical sensitivity, vomiting, and other discomforts. Anxiety and discomforts, which are experienced during labor, may be related to psychological status. Future research is needed to focus on the psychological efficacy of aromatherapy.

Conclusions

The use of aromatherapy during labor and delivery has continued to expand in practice settings. There was a low certainty of evidence of subjective labor pain reduction by inhalation aroma essence on the latent and early active phase. However, other outcomes such as pain relief on the late active phase, anxiety relief, duration of contraction, labor length, types of delivery, labor augmentation, and the Apgar score of infants could not reach the level of evidence indicating the definitive effectiveness of aromatherapy. Some discomforts during child birth are related to pain, and this feeling of pain may be reduced following aromatherapy.

List Of Abbreviations

RCTs  Randomized control trials
MD    Mean difference
SMD   Standardized mean difference
CI     Confidence interval
SD     Standard deviation
RR     Risk ratio
ITT    Intention-to-treat
GIVM   General inverse variance method
RevMan Review Manager
VAS    Visual Analogue Scale
NRS    Numerical Rating Scale
STAI  Spielberger’s State-Trait Anxiety Inventory

VASA  Visual Analog Scale for Anxiety

Declarations

Competing interests

No competing interests have been declared by the authors.

Acknowledgement

We thank Dr. Farhad Shokraneh for assistance in the study search. We also thank Dr. Edward F. Barroga and Dr. Sarah E. Porter for advice in writing the paper.

References


Tables

Due to technical limitations, tables are only available as a download in the supplemental files section.

Figures
Figure 1

Identification

362 records identified through database searching
410 records updated to identify through database searching

Screening

520 duplicates removed

254 records screened

176 records excluded

Eligibility

78 full text records assess for eligibility

67 full text records excluded because of:
- duplicates: 3
- wrong intervention: 12
- wrong population: 6
- wrong outcome: 2
- wrong study design: 11
- wrong language: 6
- unpublished: 24
- no available data: 3

Included

10 trials 11 reports included in quantitative synthesis
<table>
<thead>
<tr>
<th>Study</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
</tr>
</thead>
</table>

**Figure 2**
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental Mean Difference</th>
<th>SE</th>
<th>Control Total</th>
<th>Control Weight</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A B C D E F G</td>
</tr>
<tr>
<td>Esmaeilzadeh-Saeieh 2018</td>
<td>-1.7 0.2009</td>
<td>62</td>
<td>62 12.9%</td>
<td>-1.70 [-2.09, -1.31]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Hamdami 2018</td>
<td>-3.11 0.1945</td>
<td>55</td>
<td>55 12.9%</td>
<td>-3.11 [-3.49, -2.73]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Kaviani 2014a</td>
<td>-0.34 0.1937</td>
<td>104</td>
<td>52 12.9%</td>
<td>-0.34 [-0.72, 0.04]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Kaviani 2014b</td>
<td>-1.2 0.335</td>
<td>80</td>
<td>80 12.4%</td>
<td>-1.20 [-1.84, -0.56]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Namazi 2014b</td>
<td>-3.11 0.1336</td>
<td>57</td>
<td>56 13.1%</td>
<td>-3.11 [-3.37, -2.85]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Tanviss 2018</td>
<td>-0.72 0.4364</td>
<td>52</td>
<td>52 11.8%</td>
<td>-0.72 [-1.58, 0.14]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Vakilian 2018</td>
<td>-0.54 0.3659</td>
<td>59</td>
<td>60 12.2%</td>
<td>-0.54 [-1.26, 0.18]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Yazdkhast 2016</td>
<td>-1.6 0.4036</td>
<td>60</td>
<td>59 12.0%</td>
<td>-1.60 [-2.39, -0.81]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
</tbody>
</table>

Total (95% CI) | 529 476 100.0% | -1.56 [-2.45, -0.67] |

Heterogeneity: $I^2 = 97%$ Test for overall effect: $Z = 3.44 (P = 0.0006)$

**Risk of bias legend**

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

---

**Figure 3**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental Mean Difference</th>
<th>SE</th>
<th>Control Total</th>
<th>Control Weight</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A B C D E F G</td>
</tr>
<tr>
<td>Esmaeilzadeh-Saeieh 2018</td>
<td>-1.44 0.2426</td>
<td>62</td>
<td>62 17.1%</td>
<td>-1.44 [-1.92, -0.96]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Hamdami 2018</td>
<td>-3.31 0.1171</td>
<td>55</td>
<td>55 17.9%</td>
<td>-3.31 [-3.54, -3.08]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Namazi 2014b</td>
<td>-2.02 0.0991</td>
<td>57</td>
<td>56 17.9%</td>
<td>-2.02 [-2.21, -1.83]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Tanviss 2018</td>
<td>-0.57 0.4475</td>
<td>52</td>
<td>52 14.9%</td>
<td>-0.57 [-1.45, 0.31]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Vakilian 2018</td>
<td>-0.64 0.3514</td>
<td>59</td>
<td>60 16.0%</td>
<td>-0.64 [-1.33, 0.05]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Yazdkhast 2016</td>
<td>-1.93 0.3317</td>
<td>59</td>
<td>60 16.2%</td>
<td>-1.90 [-2.55, -1.25]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
</tbody>
</table>

Total (95% CI) | 344 345 100.0% | -1.69 [-2.50, -0.89] |

Heterogeneity: $I^2 = 96%$ Test for overall effect: $Z = 4.13 (P = 0.0001)$

**Risk of bias legend**

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

---

**Figure 4**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental Mean Difference</th>
<th>SE</th>
<th>Control Total</th>
<th>Control Weight</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Mean Difference IV, Random, 95% CI</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A B C D E F G</td>
</tr>
<tr>
<td>Esmaeilzadeh-Saeieh 2018</td>
<td>-1.36 0.2712</td>
<td>62</td>
<td>62 16.8%</td>
<td>-1.36 [-1.89, -0.83]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Hamdami 2018</td>
<td>-3.09 0.085</td>
<td>55</td>
<td>55 17.9%</td>
<td>-3.09 [-3.26, -2.92]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Namazi 2014b</td>
<td>-1.89 0.1029</td>
<td>57</td>
<td>56 17.9%</td>
<td>-1.89 [-2.09, -1.69]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Tanviss 2018</td>
<td>-0.17 0.4299</td>
<td>52</td>
<td>52 15.1%</td>
<td>-0.17 [-1.01, 0.67]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Vakilian 2018</td>
<td>-0.81 0.3671</td>
<td>59</td>
<td>60 15.8%</td>
<td>-0.81 [-1.53, -0.09]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
<tr>
<td>Yazdkhast 2016</td>
<td>-1.47 0.3066</td>
<td>59</td>
<td>60 16.4%</td>
<td>-1.47 [-2.07, -0.87]</td>
<td></td>
<td></td>
<td>7 7 7 7 7 7 7</td>
</tr>
</tbody>
</table>

Total (95% CI) | 344 345 100.0% | -1.52 [-2.33, -0.71] |

Heterogeneity: $I^2 = 97%$ Test for overall effect: $Z = 3.67 (P = 0.0002)$

**Risk of bias legend**

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

---

**Figure 5**
**Figure 6**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental Mean</th>
<th>SD</th>
<th>Total</th>
<th>Control Mean</th>
<th>SD</th>
<th>Total</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fakari 2015</td>
<td>52.73</td>
<td>11.7</td>
<td>49</td>
<td>52.31</td>
<td>6.2</td>
<td>48</td>
<td>0.04 [-0.35, 0.44]</td>
<td>-0.35 [-0.44]</td>
<td>B</td>
</tr>
<tr>
<td>Khreikhan 2014</td>
<td>4.21</td>
<td>2.31</td>
<td>36</td>
<td>5.53</td>
<td>1.98</td>
<td>36</td>
<td>-0.70 [-1.18, -0.23]</td>
<td>-1.18 [-0.23]</td>
<td>B</td>
</tr>
<tr>
<td>Namazi 2014a</td>
<td>45.32</td>
<td>1.22</td>
<td>57</td>
<td>56.38</td>
<td>1.13</td>
<td>56</td>
<td>-9.34 [-10.63, -8.05]</td>
<td>-10.63 [-8.05]</td>
<td>A</td>
</tr>
</tbody>
</table>

Total (95%) CI: 197, 195, 100.0% -3.49 [-6.28, -0.69]

Heterogeneity: $\text{I}^2 = 99\%$
Test for overall effect: $Z = 2.45$ ($P = 0.01$)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

**Figure 7**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental Mean</th>
<th>SD</th>
<th>Total</th>
<th>Control Mean</th>
<th>SD</th>
<th>Total</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
<th>Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamdamin 2018</td>
<td>55.14</td>
<td>3.42</td>
<td>55</td>
<td>75.51</td>
<td>3.55</td>
<td>55</td>
<td>-6.80 [-6.67, -4.94]</td>
<td>-6.67 [-4.94]</td>
<td>C</td>
</tr>
<tr>
<td>Khreikhan 2014</td>
<td>2.25</td>
<td>1.71</td>
<td>36</td>
<td>4.67</td>
<td>2.74</td>
<td>36</td>
<td>-1.05 [-1.54, -0.55]</td>
<td>-1.54 [-0.55]</td>
<td>A</td>
</tr>
</tbody>
</table>

Total (95%) CI: 148, 147, 100.0% -5.54 [-10.39, -0.69]

Heterogeneity: $\text{I}^2 = 99\%$
Test for overall effect: $Z = 2.24$ ($P = 0.03$)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.pdf
- Table2.pdf
- Table1.pdf