Conservative Treatment of Stress Urinary Incontinence: A Systematic Review with Meta-analysis of Randomized Controlled Trials

Tratamento conservador da incontinência urinária: revisão sistemática e metanálise de ensaios clínicos randomizados

Rafael Mendes Moroni1,2 Pedro Sergio Magnani1 Jorge Milhem Haddad3 Rodrigo de Aquino Castro4 Luiz Gustavo Oliveira Brito1

1 Department of Gynecology and Obstetrics, Ribeirão Preto School of Medicine, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil
2 Department of Obstetrics, Gynecology and Reproductive Medicine, Universidade Estadual do Oeste do Paraná, Cascavel, West Paraná, Brazil
3 Department of Gynecology, School of Medicine, Universidade de São Paulo, São Paulo, São Paulo, Brazil
4 Department of Gynecology, Universidade Federal de São Paulo, São Paulo, São Paulo, Brazil

Address for correspondence Luiz Gustavo Oliveira Brito, MD, PhD, Department of Gynecology and Obstetrics, Ribeirão Preto School of Medicine, University of São Paulo, Avenida Bandeirantes, 3900–8th floor – Monte Alegre – Ribeirão Preto, SP, Brazil (e-mail: lbrito@usp.br).


Abstract

We performed a systematic review and meta-analysis of randomized controlled trials that studied the conservative management of stress urinary incontinence (SUI). There were 1058 results after the initial searches, from which 37 studies were eligible according to previously determined inclusion criteria. For the primary outcomes, pelvic floor muscle training (PFMT) was more efficacious than no treatment in improving incontinence-specific quality of life (QoL) scales (SMD = −1.24 SDs; CI 95% = −1.77 to −0.71 SDs). However, its effect on pad tests was imprecise. Combining biofeedback with PFMT had an uncertain effect on QoL (MD = −4.4 points; CI 95% = −16.69 to 7.89 points), but better results on the pad test, although with elevated heterogeneity (MD = 0.9 g; 95% CI = 0.71 to 1.10 g); group PFMT was not less efficacious than individual treatment, and home PFMT was not consistently worse than supervised PFMT. Both intravaginal and superficial electrical stimulation (IES and SES) were better than no treatment for QoL and pad test. Vaginal cones had mixed results. The association of IES with PFMT may improve the efficacy of the latter for QoL and pad test, but the results of individual studies were not consistent. Thus, there is evidence of the use of PFMT on the treatment of SUI, with and without biofeedback.

Keywords
► urinary incontinence
► physiotherapy
► systematic review
► electrical stimulation
► meta-analysis

ISSN 0100-7203.

Copyright © 2016 by Thieme Publicações Ltda, Rio de Janeiro, Brazil

License terms
Introduction

Urinary incontinence (UI) is a very common condition among women, with reported annual incidences of 1% to 35%\(^1\), and a lifetime risk of surgery of 11% to 13%.\(^2\) Its impact on quality of life is significant, leading to physical and social limitations, shame, and increased rates of depressive symptoms.\(^4\) Stress urinary incontinence (SUI) is the most common subtype of UI, and its treatment is predominantly surgical, particularly after the introduction of synthetic midurethral slings in the mid-nineties, which led to less invasive surgical procedures. Despite its efficacy, surgical treatment may lead to complications, directly or indirectly related to the use of synthetic materials.\(^5\) Therefore, conservative treatment is considered the first-line option for managing women with SUI.\(^6\)

However, multiple alternatives are available, such as pelvic floor muscle training (PFMT), bladder training, electrical stimulation and vaginal cones, and the relative efficacy between all these options is not well documented. We have retrieved a few systematic reviews with meta-analysis which addressed the comparison between PFMT and no treatment,\(^7\) and PFMT plus an alternative treatment versus that alternative treatment alone.\(^8\) However, the objective of this review is different from the previously published systematic reviews; we aimed to pool randomized trials which compared multiple forms of conservative treatment (alone or in association) between each other, with control groups or surgical treatments, emphasizing treatment options that are available in Brazil due to the lack of guidelines for these practitioners.

Methods

We entered a protocol for this review, which is currently registered in the PROSPERO database (CRD42015020645; available at: www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015020645).

Only randomized controlled trials (RCTs) were considered eligible. We excluded other study designs (such as cohorts, case-controls, quasi-randomized trials). We have also chosen to consider a minimal methodological quality as an inclusion criteria, and have used the Jadad scale\(^9\) for classifying the RCTs. Studies that obtained a score lower than three in such scale were not considered eligible for inclusion due to a higher risk of bias.

The target population was formed by adult women, aged 18 years or older, with a clinical diagnosis of SUI (complaint, and/or an observation during examination of urinary leakage due to effort or straining), with absence of neurological injuries or diseases. Urodynamic diagnosis of SUI was not considered necessary for inclusion. Studies that comprised women with urgency urinary incontinence or mixed urinary incontinence were excluded, since these diagnoses differ significantly from SUI, and the observed response to treatment could also be significantly different.

The interventions were any forms of conservative treatment for SUI, compared against each other, either alone or in combination with one another. Studies which assessed treatments or devices unavailable in Brazil were not included.

Variables

Primary outcomes were: incontinence-related quality of life, measured through validated instruments, such as the King's Health Questionnaire (KHQ) and the Incontinence Quality of Life Questionnaire (I-QoL); objective measure of incontinence, quantified in grams through pad-tests.

Secondary outcomes: Number of incontinence episodes, measured through bladder diaries; Subjective improvement,
measured through visual scales, such as Likert scales, or through the number of satisfied patients; General quality of life, measured through non-specific instruments, such as SF-36; Adverse events.

**Study Search and Selection**

We performed searches for RCTs in the following electronic databases: Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, and LILACS. We then imported the results to EndNote reference management software (Thomson Reuters, New York, U.S.A.). We adapted the search strategy according to the rules of each database, using the following elements: (stress urinary incontinence) OR (stress incontinence) AND ((treatment) OR (therapy)) AND (conservative) OR (medical) OR (physical therapy) OR (physiotherapist) OR (electrical stimulation) OR (PFMT) OR (training) OR (estrogen)). Furthermore, we used the “Cochrane highly sensitive filter for detecting RCTs” in the MEDLINE search by adding the following elements to the search strategy: (randomized controlled trial[pt]) OR (controlled clinical trial [pt]) OR (randomized[tiab]) OR (placebo[tiab]) OR (clinical trials as topic[mesh:noexp]) OR (randomly[tiab]) OR (trial [ti]) NOT ((animals[mh] NOT (humans[mh] AND animals [mh]))).

There were no restrictions on language or year of publication; we searched all the databases from their inception up to the search date. Two study authors (RMM and LGB) performed the initial screening independently, by reading titles and abstracts and locating potentially eligible entries. Then, the same two authors obtained and analyzed independently the full text of each potentially eligible result, to determine true eligibility. Any disagreements between the two authors were resolved by consulting a third author (PSM).

**Data Collection and Extraction**

Data extracted from the publications of eligible studies were: country, institution, conflicts of interest, ethical approval, participants’ consent, data collection period, inclusion and exclusion criteria, characteristics of participants, treatments, and outcomes of interest. Two authors (RMM and LGB) performed data extraction independently by using a data extraction form developed and pilot-tested by the authors.

**Statistical Analyses**

We used a statistical package (RevMan 5.3; Nordic Cochrane Center, Copenhagen, Denmark) to perform the analysis. We summarized binary outcomes (such as subjective improvement measured through the number of satisfied patients; adverse events) using Mantel-Haenszel relative risks (RR). We preferred RR over odds ratios (OR) due to its easier interpretation. However, we have used Peto OR in situations with zero number of events in one of the groups, because it is the most accurate measure in such situations. We summarized continuous outcomes (all the other outcomes of interest – quality of life scores in multiple scales; pad test weights; number of incontinence episodes; visual analog scales) using the mean difference (MD), calculated through the inverse variance method. Precision of estimates appear as 95% confidence intervals (95%CIs). We quantified statistical heterogeneity by using $I^2$, and informed its value together with the estimates. When we measured a single continuous outcome using different scales in different studies (such as incontinence-specific quality of life measured through different instruments), we summarized the outcome by using the standardized mean difference (SMD). SMDs are not expressed in the original outcome scale, but in standard deviations (SDs), which are interpreted as follows: SMD between 0.2 and 0.5 SD indicates a small effect; between 0.5 and 0.8 SD, a moderate effect; and larger than 0.8 SD, a large effect.

Through meta-analyses, we pooled measures of single outcomes reported by different studies that addressed similar comparisons between conservative treatment methods. We expressed pooled data graphically through forest-plots, in which an increase in the measure of an outcome is shown to the right of the central line, and such an increase may be beneficial (such as an increase in a certain quality of life scale score) or harmful (such as an increase in the number of episodes of incontinence). We chose to pool the studies in which the interventions were actually comparable and in which the study groups were also comparable before the interventions. We evaluated statistical heterogeneity, expressed by the $I^2$ value for each meta-analysis; we considered it elevated when higher than 50%. In situations of elevated heterogeneity, we reevaluated the individual studies to assure that we were actually studying interventions and populations that were comparable. If they were indeed similar, we chose to perform random effects meta-analyses.

**Quality of Included Studies**

The Jadad scale was used to judge the methodological quality of individual studies and previously described. Two authors (RMM and LGB) independently assessed the methodological quality of individual studies, giving an appropriate to score each study. The two authors resolved disagreements through discussion or by consulting a third author (PSM).

**Results**

We performed electronic searches on May 10th 2015 and obtained a total of 1058 hits: 590 from MEDLINE, 364 from CENTRAL, and 104 from LILACS. Furthermore, we located one additional entry through the revision of reference lists from other similar reviews. The study selection process is illustrated in a PRISMA flowchart presented in Fig. 1. Table 1 shows the characteristics of included studies. Pooled data are represented in Figs. 2 and 3. The interventions included in this review are the following: Pelvic floor muscle training (PFMT); Biofeedback; Bladder training; Electrical stimulation; Vaginal cones; Magnetic stimulation; and Paula method.

**Pelvic Floor Muscle Training (with or without Biofeedback) versus No Treatment**

Three eligible studies studied such a comparison, which included a total of 122 patients, 59 in the PFMT groups and 66 in the control groups.
Incontinence-specific quality of life – two studies\(^{11,12}\) reported this outcome, the first one measuring it using the King’s Health Questionnaire (KHQ – higher scores are worse) and the second using IIQ-7 (higher scores are better – we inverted the direction of effect in the meta-analysis to allow combination with KHQ results). Meta-analysis (►Fig. 2A) demonstrated a moderate to large effect of PFMT toward a better score in incontinence-specific scales.

Pad-test (g) – one study\(^ {12}\) reported that there was no sufficient precision in the study to distinguish a beneficial outcome from a harmful effect of PFMT.

- Number of incontinence episodes – one study\(^ {12}\) demonstrated a beneficial effect of PFMT: a decrease in the number of incontinence episodes.
- General quality of life – one study measured this outcome\(^ {13}\) through the Quality of Life Scale (QoLS), and although there was a tendency toward a beneficial effect, the result was not sufficiently precise to distinguish a harmful, null, or beneficial effect of PFMT.

**PFMT versus PFMT Plus Biofeedback**

Five studies were eligible\(^ {14-18}\), which included 250 women, 126 in the PFMT groups and 124 in the PFMT, in addition to biofeedback groups. One of them\(^ {15}\) was not included in the meta-analyses for some of the outcomes because its groups were heterogeneous at the beginning of the study, and the use of final measures only in the pooled estimates would not take this initial difference between groups into consideration.

- Incontinence-specific quality of life – one study measured this with the KHQ.\(^ {17}\) Considering the domain *incontinence impact*, the study was not sufficiently precise to distinguish a beneficial, a harmful or a null effect of the association of PFMT with biofeedback, when compared with PFMT alone.
- Pad test (g) – A meta-analysis of four studies\(^ {14,16-18}\) (►Fig. 3A) resulted in a sufficiently precise estimate demonstrating higher urine loss in the PFMT groups, when compared with PFMT + biofeedback, suggesting a better effect from the combined treatment. Although this result was statistically significant, it does not seem to have a large clinical relevance, since the increase in losses documented through the pad test were only around 1 g more in the PFMT group. Also, such result is due, primarily, to the result of a single study.\(^ {14}\) If such a study is excluded from the meta-analysis, the estimate from the three other studies becomes imprecise.
- Number of incontinence episodes – one study calculated this,\(^ {17}\) but it was not precise enough to distinguish a better, a worse or an indifferent effect when comparing PFMT + biofeedback with PFMT alone.
- Subjective improvement – two studies reported this outcome,\(^ {14,18}\) the first one using a visual analog scale (0 to 10) and the second one using a Likert scale (0 to 4). The combination of their results was not sufficiently precise to demonstrate whether combined treatment was better, worse, or indifferent when compared to PFMT alone.
- Adverse events – three studies reported adverse events.\(^ {15,17,18}\) There was not enough precision to distinguish any difference from a larger number of adverse events in any of the groups when comparing PFMT alone with combined treatment.
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Treatments compared</th>
<th>Treatment duration</th>
<th>Patients analyzed per group</th>
<th>Diagnosis of incontinence</th>
<th>Jadad scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahlund et al</td>
<td>Sweden</td>
<td>Postpartum supervised PFMT vs written instructions for PFMT postpartum</td>
<td>24 wk.</td>
<td>40 / 42</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Aksac et al</td>
<td>Turkey</td>
<td>Home PFMT vs Home PFMT + BF vs Control</td>
<td>8 wk.</td>
<td>20 / 20 / 10</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Auhee et al</td>
<td>Finland</td>
<td>Home PFMT vs Home PFMT + BF</td>
<td>12 wk.</td>
<td>15 / 15</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Berghmans et al</td>
<td>Holland</td>
<td>Supervised PFMT vs Supervised PFMT + BF</td>
<td>4 wk.</td>
<td>20 / 20</td>
<td>Clinical and/or UD</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Bo et al</td>
<td>Norway</td>
<td>Home PFMT (12 contr. 3x/day) vs Supervised (45min/wk.) and home PFMT</td>
<td>24 wk.</td>
<td>29 / 23</td>
<td>UD and pad test</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Bo et al</td>
<td>Norway</td>
<td>Supervised PFMT vs IES (50Hz; 30min/day) vs Cones (20min/day) vs Control</td>
<td>24 wk.</td>
<td>25 / 25 / 27 / 30</td>
<td>UD and pad test</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Bo et al</td>
<td>Norway</td>
<td>Home and supervised PFMT vs Control</td>
<td>24 wk.</td>
<td>25 / 30</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Borello et al</td>
<td>USA</td>
<td>Supine PFMT vs Supine and standing PFMT</td>
<td>9 to 12 wk.</td>
<td>17 / 19</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Cammu et al</td>
<td>Belgium</td>
<td>PFMT + BF (30min 1x/wk.) vs Cones (15min. 2x/day)</td>
<td>12 wk.</td>
<td>30 / 30</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Castro et al</td>
<td>Brazil</td>
<td>Supervised PFMT vs IES(50Hz; 20min/day) vs Cones (25min 3x/wk.) vs Control</td>
<td>24 wk.</td>
<td>26 / 27 / 24 / 24</td>
<td>UD</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Correia et al</td>
<td>Brazil</td>
<td>SES (50Hz; 20min; 2x/wk.) vs IES (50Hz; 20min; 2x/wk.) vs Control</td>
<td>6 wk.</td>
<td>15 / 15 / 15</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>De Oliveira et al</td>
<td>Brazil</td>
<td>Group PFMT (45min, 2x/wk) vs Individual supervised PFMT (30min, 2x/wk)</td>
<td>12 wk.</td>
<td>30 / 30</td>
<td>UD</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Dumoulin et al</td>
<td>Canada</td>
<td>Supervised PFMT vs Supervised PFMT + abdominal training vs Control</td>
<td>8 wk.</td>
<td>20 / 23 / 19</td>
<td>UD and pad test</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Felicissimo et al</td>
<td>Brazil</td>
<td>Home PFMT (90 a 180 contr./day) vs Supervised PFMT (50min, 2x/wk) + Home PFMT</td>
<td>8 wk.</td>
<td>30 / 29</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Treatments compared</td>
<td>Treatment duration</td>
<td>Patients analyzed per group</td>
<td>Diagnosis of incontinence</td>
<td>Jadad scale</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>----------------------------------------------------------</td>
<td>--------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Ferreira et al&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Portugal</td>
<td>Home PFMT (30 contr./day) vs Supervised (45min. 1x/wk) + home PFMT</td>
<td>24 wk.</td>
<td>17 / 17</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Fitz et al&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Brazil</td>
<td>PFMT + BF (40min 2x/wk) vs Control</td>
<td>4 wk.</td>
<td>16 / 16</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Gilling et al&lt;sup&gt;50&lt;/sup&gt;</td>
<td>New Zealand</td>
<td>Magnetic stimulation (10-50Hz, 20min, 3x/wk) vs Placebo/Sham</td>
<td>6 wk.</td>
<td>35 / 35</td>
<td>UD</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Harvey et al&lt;sup&gt;43&lt;/sup&gt;</td>
<td>Canada</td>
<td>Supervised + home PFMT vs Cones</td>
<td>6 wk.</td>
<td>7 / 7</td>
<td>UD</td>
<td>3 (2 / 1 / 0)</td>
</tr>
<tr>
<td>Hirakawa et al&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Japan</td>
<td>Home PFMT (2x/day) vs Home PFMT + BF (2x/day)</td>
<td>12 wk.</td>
<td>23 / 23</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Kamel et al&lt;sup&gt;45&lt;/sup&gt;</td>
<td>Egypt</td>
<td>Abdominal training (3x/wk) vs Supervised PFMT (3x/wk)</td>
<td>12 wk.</td>
<td>15 / 15</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Kaya et al&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Turkey</td>
<td>Bladder training vs PFMT + bladder training</td>
<td>6 wk.</td>
<td>24 / 26</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Knight et al&lt;sup&gt;31&lt;/sup&gt;</td>
<td>UK</td>
<td>Supervised IES (35Hz; 35 min; 1x/10days) vs Home IES (10Hz, daily during night sleep) vs Home PFMT (6x/day)</td>
<td>24 wk.</td>
<td>19 / 20 / 18</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Liebergall et al&lt;sup&gt;49&lt;/sup&gt;</td>
<td>Israel</td>
<td>Paula Method (45min/day) vs Group + home PFMT</td>
<td>12 wk.</td>
<td>117 / 123</td>
<td>Pad test</td>
<td>3 (2 / 1 / 0)</td>
</tr>
<tr>
<td>Luber et al&lt;sup&gt;34&lt;/sup&gt;</td>
<td>USA</td>
<td>IES (50Hz; 15min, 2x/day) vs Placebo/Sham</td>
<td>12 wk.</td>
<td>20 / 24</td>
<td>UD</td>
<td>3 (2 / 1 / 0)</td>
</tr>
<tr>
<td>McLean et al&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Canada</td>
<td>Supervised + domestic PFMT vs Control</td>
<td>12 wk.</td>
<td>18 / 17</td>
<td>UD</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Miller et al&lt;sup&gt;27&lt;/sup&gt;</td>
<td>USA</td>
<td>Pre-contraction (&quot;Knack&quot;) vs Control</td>
<td>1 wk.</td>
<td>13 / 14</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
<tr>
<td>Morkved et al&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Norway</td>
<td>Supervised and domestic PFMT vs Supervised and domestic PFMT + BF</td>
<td>24 wk.</td>
<td>48 / 46</td>
<td>UD and pad test</td>
<td>4 (2 / 1 / 1)</td>
</tr>
<tr>
<td>Parsons et al&lt;sup&gt;32&lt;/sup&gt;</td>
<td>UK</td>
<td>PFMT + IES vs PFMT + placebo vs PFMT + control</td>
<td>14 wk.</td>
<td>82 / 42 / 40 / 20</td>
<td>UD</td>
<td>3 (1 / 1 / 1)</td>
</tr>
<tr>
<td>Pereira et al&lt;sup&gt;22&lt;/sup&gt;</td>
<td>Brazil</td>
<td>Group PFMT (2x/wk, 60 min) vs</td>
<td>12 wk.</td>
<td>15 / 15 / 15</td>
<td>Clinical</td>
<td>3 (2 / 0 / 1)</td>
</tr>
</tbody>
</table>
Beyond these five mentioned studies, a sixth study assessed the use of biofeedback, but made a somewhat different comparison, and thus was not included in the quantitative analyses.

**Supine and Standing PFMT versus Supine Only PFMT**
This comparison was performed by one eligible study, which included 36 women, 19 in the supine and standing PFMT group, and 17 in the supine only PFMT group.
- Incontinence-specific quality of life – this outcome was assessed using the IIQ scale (higher scores are worse). The result was not sufficiently precise to distinguish a better, worse, or indifferent effect when combining both positions, compared with supine only PFMT.
- Pad test (g) – Likewise, the result was imprecise to detect any significant difference between the combination of positions and supine only PFMT.

**Individual PFMT versus Group PFMT**
Two studies were eligible. They included 90 women, 45 in the individual PFMT groups and 45 in the PFMT groups. We included both studies in the quantitative analyses for at least one outcome.
- Incontinence-specific quality of life – both studies assessed this outcome using KHQ (higher scores are worse). Meta-analysis (► Fig. 2B) resulted in an imprecise estimate to determine whether individual PFMT is better, worse, or indifferent when compared to the PFMT group for this outcome.
• Pad test (g) – both studies showed that the pooled estimate (Fig. 3B) was not precise enough to determine whether individual PFMT was superior or indifferent when compared with the PFMT group.

• Number of incontinence episodes – only one study reported this outcome and its estimate was not precise enough to detect any difference between individual and the group PFMT.

• Subjective improvement (number of satisfied women) – both studies reported this outcome. Meta-analysis resulted in an imprecise estimate to distinguish a superior effect (higher number of satisfied women), from an inferior effect (lower number of satisfied women) or from no difference when comparing individual PFMT with group PFMT.

**Home PFMT versus Intensive PFMT with Physiotherapist**

Four studies assessed this comparison, including a total of 189 women, 97 in the home PFMT groups and 92 in the intensive PFMT groups. All four studies were included in the quantitative analyses for at least one outcome.

• Incontinence-specific quality of life – three studies reported this outcome, but the combination of individual results through meta-analysis was not possible. One of them assessed the outcome using the Ditrovie scale (higher scores represent a worse quality of life). The study estimate was not sufficiently precise to determine if intensive PFMT was more effective, less effective or no different from home PFMT. Another study used the ICIQ-SF scale and demonstrated an equivalent improvement in both groups (final median in the intensive group = 8 points (5 to 13) / in the home PFMT group = 8 points (6 to 12); p = 0.7). The third study, on the other hand, demonstrated better results in the intensive PFMT group using the I-QoL scale for assessment, in which higher scores represent a better quality of life (median in the intensive group = 89 points / in the home PFMT group = 79 points; p = 0.04).

• Pad test – two studies assessed this outcome, both reporting it with medians only, which did not allow the calculation of combined estimates. Individual study results are, thus, reported. In the first study, the final median in the intensive group was 3.2 g (1.2 g to 8 g), while in the home PFMT was 2.8 g (1.5 g to 8.5 g); p = 0.78; there was no significant difference between the groups. In the second study, the final median in the intensive group was 3.2 g, while in the home PFMT it was 15 g; p = 0.0018; the intensive group had significantly less urine loss on the pad test.

• Number of incontinence episodes – one study reported that this estimate was not sufficiently precise to determine if intensive PFMT was more efficacious, less efficacious or no different from home PFMT.

• Subjective improvement (number of women with significant improvement or cure) – three studies reported this outcome. A pooled estimate from all three studies was precise enough to determine an inferior effect of home PFMT when compared with intensive PFMT, with a significantly inferior number of satisfied women in the home PFMT groups. A high heterogeneity was observed in this meta-analysis, which was related to a single study, which, differently from the other two studies, did not find any significant difference between the groups. If this single study is excluded, the heterogeneity is solved, and the estimate still demonstrates a better effect in the intensive PFMT group.

**Use of a Pre-contraction (Knack) versus No Maneuver**

A single study compared the use of a pre-contraction right before a physical effort, versus the same effort without a pre-contraction, including 27 women, 13 in the pre-contraction group, and 14 in the no maneuver group.

• Pad test – only this outcome was studied for this comparison. The authors performed a modified test, in which a paper towel was used, and the area of the spot resulting from wetting the towel was measured in cm². There was a strong tendency of the pre-contraction to be associated with a smaller wetted area, but the estimate was imprecise to distinguish a beneficial, a harmful or a null effect from performing a pre contraction.

**PFMT versus Bladder Training (with or without Associated PFMT)**

This comparison was assessed by two studies, which included a total of 126 women, 61 in the PFMT group and 65 in the bladder training group. The results of these studies were not combined, because one of them reported its results using only medians. Individual study results are thus reported, here, for each informed outcome.

• Incontinence-specific quality of life – one study assessed this outcome using the ICIQ-SF scale (higher scores indicate worse quality of life) and the estimate was precise enough to indicate a higher quality of life after treatment with PFMT. Other studies used the I-QoL scale (higher scores indicate worse quality of life) – there was a median decrease in IQ-7 score to 19 (9.5 to 42.8) in the PFMT group, and equal to 7.1 (0 to 28.5) in the bladder training group; there was higher quality of life in the PFMT group at the end of the study.

• Pad test – one study reported final median values with a 95% CI demonstrating significantly lower urine loss in the PFMT group (median in the PFMT group = 0.1 g; 95% CI = 0.3 g to 3.1 g; 41 women / median in the bladder training group = 0.5 g; 95% CI 1 g to 3.4 g; p = 0.03).

• Number of episodes of incontinence – one study had an estimate which was not sufficiently precise to distinguish a larger, lower or indifferent number of episodes of loss in the PFMT group, when compared with the bladder training group.

• Subjective improvement (visual analog scale for bother – 0 to 10) – one study presented an estimate which precisely demonstrated better results in the PFMT group, with higher scores in the bother scale observed in the bladder training group.
Fig. 2  Meta-analyses for the outcome “Incontinence-specific quality of life. (A) PFMT vs control; (B) Group PFMT vs individual PFMT; (C) Intravaginal electrical stimulation vs control; (D) Superficial electrical stimulation vs control; (E) Vaginal cones vs control; (F) PFMT vs vaginal cones; (G) Intravaginal electrical stimulation vs vaginal cones. Abbreviations: SD, standard deviation; IV, inverse variance; IES, intravaginal electrical stimulation; SES, superficial electrical stimulation.
Fig. 3  Meta-analyses for the outcome “Pad test” (A) PFMT vs PFMT + biofeedback; (B) Group PFMT vs individual PFMT; (C) Intravaginal electrical stimulation vs control; (D) Superficial electrical stimulation vs control; (E) Vaginal cones vs control; (F) PFMT vs vaginal cones; (G) Intravaginal electrical stimulation vs vaginal cones. Abbreviations: SD, standard deviation; IV, inverse variance; IES, intravaginal electrical stimulation; SES, superficial electrical stimulation.
• General quality of life – one study\(^\text{30}\) used the AQoL scale (higher scores indicate better quality of life), and there was not enough precision to distinguish a beneficial, harmful or indifferent effect when comparing bladder training to PFMT.

**PFMT Plus Intravaginal Electrical Stimulation (IES) versus PFMT Alone**
This comparison was assessed by three studies\(^\text{31–33}\), including a total of 253 women, 166 in the PFMT + IES group and 87 in the PFMT group. One of these studies\(^\text{34}\) did not report detailed quantitative results, thus the estimates of the other two studies are reported for all given outcomes.

• Incontinence-specific quality of life – one study\(^\text{33}\) used the I-QoL scale (higher scores indicate better quality of life) and there was a precise estimate indicating better results in the PFMT + IES.

• Pad test – one study\(^\text{33}\) demonstrated a sufficiently precise estimate suggesting lower urine loss in the PFMT + IES group. Another study\(^\text{31}\) reported a median loss of 0.8 g (0 to 88 g) in the PFMT group and of 1.5 g (0 to 28 g) in the PFMT + IES group, with significant differences between the two groups.

• Number of episodes of incontinence – one study\(^\text{13}\) demonstrated a significantly lower number of episodes in the PFMT + IES group.

• Subjective improvement (number of satisfied women) – one study\(^\text{31}\) demonstrated a tendency in the PFMT + IES group to have a higher proportion of satisfied women, but the estimate was not sufficiently precise to be significant. One of the three studies assessing this outcome did not report quantitative data,\(^\text{32}\) but it concluded that there was no significant difference in the pad test or KHQ measures when combining IES with PFMT. However, in this study, we performed IES at home, and it could have suffered a lower adherence.

**Intravaginal Electrical Stimulation (IES) versus Control**
This comparison was reported by five studies\(^\text{34–39}\), including 221 women, 115 in the IES groups and 106 in the control groups.

• Incontinence-specific quality of life – this outcome was reported by two studies\(^\text{40–42}\), which have used the KHQ and I-QoL scales, respectively. Meta-analysis (\(\text{Fig. 2C}\)) of the estimates has shown a large effect of IES toward a better score in the quality of life scales, when compared with the control groups.

• Pad test – reported by four studies.\(^\text{35–39}\) The pooled estimate (\(\text{Fig. 3C}\)) demonstrated a significantly lower urine loss in the IES group, although with elevated statistical heterogeneity. The heterogeneity was related to one of the studies\(^\text{35}\), which reported the most imprecise estimate and which did not detect a difference between the groups.

• Number of incontinence episodes – reported by four studies\(^\text{34–36,39}\), with three of them included in a meta-analysis. A pooled estimate was sufficiently precise to indicate a beneficial effect of IES, when compared with the control groups. Again, one study\(^\text{35}\) did not detect a significant difference between the two groups, and such study is the main source of heterogeneity in this meta-analysis. The fourth study\(^\text{34}\) not included in the pooled estimate for not reporting detailed quantitative results, did not detect a difference between IES and the use of sham vaginal probe.

• Subjective improvement – two studies\(^\text{34,36}\) reported this outcome, measuring it by the number of satisfied women, and they did not detect a significant difference between the groups. Another two studies\(^\text{35,39}\) measured this outcome using a bother visual analog scale, but their results were not pooled due to very high heterogeneity. They both demonstrated lower bother scores in the IES groups.

• Adverse events – reported by two studies,\(^\text{35,39}\) there was a tendency towards a higher number of adverse events in the IES groups, but the estimate was not sufficiently precise.

**Surface Electrical Stimulation (SES) versus Control**
Two studies including a total of 44 women, 22 in each group, assessed this comparison.\(^\text{38,40}\)

• Incontinence-specific quality of life – a pooled estimate from both studies (\(\text{Fig. 2D}\)), both also using the KHQ scale, evidenced a better result in the SES group, when compared with the control group.

• Pad test – meta-analysis combining the results of both studies (\(\text{Fig. 2E}\)) demonstrated a precise estimate suggesting lower losses at the pad test in groups treated with SES.

**Vaginal Cones versus Control**
Such comparison was assessed by three studies\(^\text{35,36,41}\), including a total of 135 women, 66 in the vaginal cone groups and 69 in the control groups.

• Incontinence-specific quality of life – this outcome was reported by two studies\(^\text{36,41}\), which used the KHQ and the I-QoL scales, respectively. Pooled estimate from both studies (\(\text{Fig. 2E}\)) was sufficiently precise to demonstrate that the use of vaginal cones results in better quality of life scores.

• Pad test – all three studies reported this outcome, and a combined estimate was not precise enough (\(\text{Fig. 3E}\)) to distinguish no effect from a beneficial effect after using vaginal cones, when compared with the control groups.

• Number of incontinence episodes – reported by two studies\(^\text{35,36}\), a combination of their results, however, was not precise enough to distinguish if there was a beneficial effect, a harmful one or no effect when using vaginal cones, compared with the control groups. Individually, one of the studies\(^\text{35}\) did not demonstrate a significant difference between groups (95%CI = \(-1.56\) to 2.56 episodes), while the other\(^\text{36}\) demonstrated better results after using vaginal cones (95%CI = \(-9.92\) to \(-4.98\) episodes).

• Subjective improvement (number of satisfied women) – reported by only one study\(^\text{36}\), there was evidence of better
results when using vaginal cones, with a larger number of satisfied women.

**PFMT versus Vaginal Cones**

Five studies assessed this comparison, including 204 women, 101 in the PFMT groups and 103 in the vaginal cones groups.

- Incontinence-specific quality of life – three studies reported this outcome, two of them quantitatively. Meta-analysis of the results from these two last studies was imprecise to distinguish whether using PFMT had a superior, inferior or indifferent effect when compared with the use of vaginal cones. The third study did not inform quantitative results, but also concluded that there was no significant difference between the groups.

- Pad test – reported by four studies, meta-analysis of the results for this outcome was not sufficiently precise to determine whether using PFMT had a better, worse or indifferent result when compared with vaginal cones.

- Number of incontinence episodes – reported by three studies, whose combined estimates was not precise enough to determine whether the use of PFMT has a superior, inferior or indifferent effect when compared with vaginal cones.

- Subjective improvement (number of women with significant improvement). Two studies reported there was not enough precision in the pooled estimate to distinguish a better, worse, or indifferent result for this outcome after using PFMT, when compared with vaginal cones.

- Adverse events – two studies reported adverse events with a pooled estimate indicating a significantly lower number of these events in the PFMT groups.

**Intravaginal Electrical Stimulation (IES) versus Vaginal Cones**

Three studies assessed this comparison, including a total of 148 women, 76 in the IES groups and 72 in the vaginal cones groups.

- Incontinence-specific quality of life – two studies reported this outcome, both using the I-QoL scale. A combined estimate was not sufficiently precise to distinguish a beneficial, harmful or indifferent effect of IES when compared with vaginal cones.

- Pad test – three studies reported this outcome, and the combination of their results was not precise enough to determine whether IES had a superior, inferior or indifferent effect when compared with vaginal cones.

- Subjective improvement (number of patients with significant improvement) – outcome reported by two studies, whose combined estimates was not able to precisely determine whether IES was better than, worse than or equal to vaginal cones.

- Adverse events – only one study reported this outcome, and its estimate was imprecise to demonstrate any significant difference between the groups.

**PFMT versus Abdominal Training**

Two studies evaluated the training of the transverse abdominal muscles, one in an isolated fashion, and the other in combination with PFMT, and both comparing it with PFMT alone. The first study did not report any of the outcomes of interest for this review (it only reported urodynamic measures and vaginal pressure measures) but it concluded that abdominal training was more effective than PFMT in relation to these outcomes. The study was limited by the fact that it only included obese women with SUI. The second study reported its quantitative results in medians only, but concluded that the association of abdominal training to PFMT did not have any additional benefit when compared with PFMT alone (women with pad test < 2 g after treatment = 70% in the PFMT group and 73% in the abdominal + PFMT group).

**Specifically Oriented PFMT for Postpartum SUI versus General Theoretical Postpartum Instructions**

One study compared the use of an initial supervised training followed by home PFMT, with reassessments by a physical therapist every six weeks, versus the use of general theoretical postpartum instructions for PFMT, without practical orientation and without the prescription of a specific program. The study did not report any of the outcomes of interest of this review, but was limited to reporting outcomes associated with muscle strength (pemometry Oxford scale). The authors concluded that both groups had equivalent improvements, with no significant difference between them.

**PFMT versus the Paula Method**

Only one study assessed this comparison, including 176 women, 84 in the PFMT group and 92 in the Paula method group.

- Incontinence-specific quality of life – the study had a statistically precise estimate suggesting better quality of life scores with the Paula method. Clinically, however, this result may translate into an effect that goes from no clinically significant difference between the treatments up to better results with the Paula method.

- Pad test – the study estimate was not precise to distinguish if the PFMT had a better, worse or equal effect compared with the Paula method.

**Magnetic Stimulation (MS) with PFMT versus PFMT Alone**

Only one study made this comparison, including 70 women, 35 in the MS + PFMT group and 35 in the PFMT group.

- Incontinence-specific quality of life – the study used the I-QoL scale and there was insufficient precision in the estimate to determine whether the association of MS with PFMT is better, worse or no different, when compared with PFMT alone.

- Pad test – the study estimate was also imprecise to distinguish if MS + PFMT was better than or no different from PFMT alone.
Discussion

In this systematic review with meta-analysis of eligible studies, we tried to perform a very broad search, with the objective of locating any comparisons that involved the use of conservative treatments for SUI. We set a minimal methodological quality as the inclusion criterion, due to a very large predicted number of results before starting the searches. We were able to pool the results of individual studies into combined estimates using meta-analyses for various comparisons and outcomes of interest, generating estimates that comprise a larger number of participants when compared to the estimates of the original studies. Although we made an effort to pool only effectively similar comparisons, we encountered a great deal of variability in the characteristics of interventions along different studies, even when studies were assessing a same kind of intervention (e.g., studies assessing PFMT, in which one study uses an exercise program executed multiple times a day, while another study uses a similar program, but executed only once a day). Due to this variability, we observed significant statistical heterogeneity in some comparisons and outcomes. Acknowledging such potential source for heterogeneity, when performing meta-analyses in such cases, we used random-effect meta-analyses, which takes into account this variability among studies. Various studies did not report their outcomes in a way that could be used to perform meta-analyses, and these cases, we tried to report their individual results in a descriptive way.

The results of this review demonstrate that PFMT exercises are efficacious for treating SUI, being better than no treatment and resulting in higher quality of life indexes. Such observation agrees with other systematic reviews that studied this theme, but attention should paid to the fact that most studies which used urodynamic evaluations when screening potential participants excluded women with evidence of sphincter deficiency. This was probably due to the fact that these women have a smaller potential of improvement with conservative treatments. In regards to the specific approaches to performing such treatment, pooled estimates from the included studies did not demonstrate significantly different results for the studied outcomes when performing PFMT individually or in groups. Such observation is important, because group training can potentially reduce costs and increase availability of such therapy. We could not analyze characteristics such as number of repetitions of number of sessions, because variability was very large among studies. Regarding the position in which the exercises are performed, there was no significant difference when performing them in supine position or when combining both supine and standing positions. Such observation differed from the initial hypothesis in the study that assessed this comparison, since the authors expected that training performed in conditions that simulated daily activities (and thus, performed while standing) would result in superior outcomes. Exercising in a more intensive way, supervised by a physical therapist, when compared with home exercises, had some controversial results. Unfortunately, it was not possible to pool the studies’ results for some outcomes, and individual studies differed in their conclusions. Considering individual findings, there was no difference between these two modalities in some studies, while supervised PFMT was better in others. The only possible combination was for the number of women satisfied with their treatment, in which supervised PFMT was better than home exercises, but there was high statistical heterogeneity, which limits this conclusion. It is possible that such heterogeneity may be related to the individual patients’ abilities to perform their exercises at home, which may have differed among the studies; an initial supervised training session, which teaches women how to perform their exercises at home, may also have great influence in the adequacy of the exercises performed later with no supervision, and such aspect may have been different among the studies. The combination of biofeedback techniques during PFMT exercises seems not to lead to systematically better results when compared with PFMT alone in women that can adequately contract their pelvic floors, considering the included studies. Such combination, however, may be an option in women that cannot adequately isolate and contract such muscles. In these women, any additional benefit seems to be concentrated in the first weeks of treatment. Although the benefits of PFMT exercises usually take some time to become clinically significant, training aimed specifically at performing a pre-contraction (knack) of the pelvic floor immediately before straining may be efficacious in decreasing urine loss rather quickly, although the difference was not statistically significant in the study that assessed such intervention.

In regards to other techniques that potentiate or replace PFMT, IES was more efficacious than no treatment, resulting in better quality of life and in less urine loss observed at pad tests, according to combined estimates from eligible studies. There was high heterogeneity in the combination of these studies, probably related to specific differences in the treatments’ parameters and intensities among different studies, although they all used high frequency stimulation (around 50Hz) and only differed in treatment durations and pulse intervals. There was also a smaller number of incontinence episodes in the combined estimate, but the subjective improvement results were controversial among different studies.

As an alternative to IES, SES was also more efficacious than no treatment, resulting in better quality of life and in smaller urine loss at pad tests. Such result should be paid attention to, since the use of surface electrodes may be an alternative for women who do not consent with, or who experience some form of bother or pain with the use intravaginal electrodes, and for women with vaginal stenosis. The combination of PFMT with IES, when compared with PFMT alone, could be evaluated through pooled estimates, and individual studies presented controversial results for all outcomes. A possible explanation, again, is variability in the characteristics of the treatments among studies, with one of them using the intervention twice a day, while another using it every ten days. We cannot conclude, considering the data included in the review, that combining IES with PFMT is superior to PFMT alone – only one out of three
studies suggested that the combination leads to better results. In the restricted group of postpartum women with SUI, the use of PFMT techniques appears to be efficacious in increasing pelvic floor muscle strength, but there was no difference in performing these exercises at home, individually, or in a supervised fashion.

Vaginal cones also had controversial results when compared with no treatment, with combined estimates suggesting better quality of life, and a single study suggesting higher subjective improvement after using vaginal cones. On the other hand, urine loss at pad tests and number of incontinence episodes were not significantly better with vaginal cones, when compared with no treatment. Cones, due to their ease of use, may be an alternative for women that cannot adequately contract their pelvic floors for performing PFMT, but their results are less documented. When comparing PFMT and vaginal cones head to head, and also when comparing IES and vaginal cones, it was not possible to determine that any of them is superior to the other, due to lack of significant differences in any studied outcome. Magnetic stimulation and the Paula method have very little availability in Brazil, and have a potentially higher cost when compared with the other studied interventions. Also, observed results do not suggested better results for these treatments when compared with PFMT.

**Conclusion**

This systematic review is endorsed by the Urogynecology Committee of the Federação das Associações Brasileiras de Ginecologia e Obstetrícia (Brazilian Federation of the Societies of Gynecology and Obstetrics, FEBRASGO) and suggests that pelvic floor muscle therapy and behavioral therapy should be the first-line treatment for stress urinary incontinence and subsequent physical therapy techniques (such as biofeedback, vaginal cones, electrostimulation) are especially indicated when women are not able to voluntarily contract the pelvic floor muscles.

**References**