INTRODUCTION

In 1948, Kegel1 was the first to report pelvic floor muscle training (PFMT) to be effective in treatment of female urinary incontinence (UI). In spite of reports of cure rates of >84% in his series of patients, surgery soon became the first choice of treatment. Not until 1980s, there was renewed interest for conservative treatment. Today, there are >60 randomized controlled trials reporting statistically and clinically significant effects of PFMT on stress urinary incontinence (SUI) and mixed urinary incontinence (MUI) with predominately SUI symptoms, and several consensus statements based on systematic reviews have recommended conservative treatment and especially PFMT as the first choice of treatment for SUI/MUI.2–7

Subjective cure/improvement rates of PFMT reported in RCTs in studies including groups with SUI and MUI vary between 56% and 70%.7,6,9–11 Short-term (immediately after cessation of training) cure rates of 44–80%, defined as ≤2 g of leakage on different pad tests, have been found after PFMT.6,7 The highest cure rates at short-term were shown in single blind RCTs of high methodological and interventional quality.12–15 The participants had thorough individual instruction by a trained physiotherapist, combined training with biofeedback or electrical stimulation, and had close follow-up once or every second week during the intervention period. Adherence was high, and dropout was low.14–16 Since biofeedback and electrical stimulation have not been conclusively shown to give additional effect to PFMT in RCTs and systematic reviews,3–5,7,9 one could hypothesize that the key factors for success include close follow-up and high adherence to the training protocol.

While there is Level 1, grade A evidence of short-term effect of PFMT for female SUI or MUI with predominately SUI symptoms, there are still questions on the long-term outcome. In a Cochrane review evaluating PFMT versus no treatment, or inactive control treatments for UI in women, it was concluded that few data are available from long-term follow-up after cessation of supervised training.3,8 The aim of the present systematic review was to present long-term results of PFMT with or without biofeedback on SUI and MUI with predominately SUI symptoms, including both RCTs and pre-post-evaluation studies.

MATERIALS AND METHODS

Results from intervention studies with a pre- and post-test design, non-randomized controlled trials and RCTs using PFMT with or without biofeedback to treat SUI and MUI with predominately SUI symptoms are reported. Computerized search on the PubMed with the following search strategy was undertaken: Pelvic floor AND (training OR exercise OR physical activity) AND (follow-up OR long-term).

Aims: There is level 1, grade A evidence that pelvic floor muscle training (PFMT) is effective in treatment of stress urinary incontinence (SUI), but long-term outcome has been questioned. The aim of this systematic review was to evaluate the long-term outcome of PFMT for female SUI. Methods: Computerized search on PubMed up to year 2012 was undertaken with the search strategy: pelvic floor AND (urinary incontinence OR stress urinary incontinence) AND (training OR exercise OR physical activity) AND (follow-up OR long-term). Limitations were: humans, female, clinical trial, English, and adults. Inclusion criteria were: studies on SUI using PFMT with or without biofeedback as the intervention, follow-up period of ≥1 year. Exclusion criteria were studies using electrical stimulation alone and studies in the peripartum period. Results: Nineteen studies were included (1,141 women followed between 1 and 15 years). Statistical meta-analysis was not performed due to high heterogeneity. Only two studies provided follow-up interventions. Losses to follow-up during the long-term period ranged between 0% and 39%. Long-term adherence to PFMT varied between 10% and 70%. Five studies reported that the initial success rate on SUI and MUI was maintained at long-term. Long-term success based on responders to the original trial varied between 41% and 85%. Surgery rates at long term varied between 4.9% and 58%. Conclusions: Short-term outcome of PFMT can be maintained at long-term follow-up without incentives for continued training, but there is a high heterogeneity in both interventional and methodological quality in short-and long-term pelvic floor muscle training studies. Neurourol. Urodynam. 32:215–223, 2013. © 2012 Wiley Periodicals, Inc.

Key words: exercise; follow-up; pelvic floor; urinary incontinence
activity) AND (urinary incontinence OR stress urinary incontinence) AND (follow-up OR long-term) with the following limits: adults, female, clinical trial, English, and all adults. In addition, computerized search on the PEDro database, abstracts from the International Continence Society (ICS) and International Association of Urogynecology (IUGA) from 1990 onwards, and hand-searching of reference list of studies eligible for inclusion and former systematic reviews and guidelines were carried out.2-7,17

Long-term was defined as ≥1 year follow-up time after cessation of the original PFMT intervention. Excluded were studies in the peripartum period and studies using electrical stimulation only. Two researchers extracted data from the studies and classified them independently. Each study was classified according to pre-set criteria: original design, original intervention, short-term effect, length of the long-term follow-up period, whether there was follow-up intervention (yes or no), description of outcome measure at long-term follow-up, loss to follow-up and adherence to PFMT in the follow-up period and long-term outcome. Surgery rate during the follow-up period was the pre-set primary outcome and report of cure/maintenance of improvement was the secondary outcome. The PRISMA statement for reporting systematic reviews was followed.18,19 For controlled studies, scores of internal validity given by independent raters of the PEDro database were used if available, if not, they were scored independently by the two reviewers using the PEDro score.20 PEDro is a 10 point scale giving 1 point for each of the following criteria: random allocation, concealed allocation, baseline comparability, blinding of subjects, blinding of therapist, blinding of assessor, adequate follow-up (>85%), intention to treat (ITT) analysis, report of statistical comparison between groups and provision of point estimates and measures of variability.

RESULTS

Search on PubMed identified 44 studies, with 17 long-term studies of PFMT fulfilling the inclusion criteria. Two additional studies were found by hand search of reference lists. The 19 studies included 1,141 women and are presented in Table 1.11,20-37 Three research groups reported long-term results for the same original study at two time points (21 and 30, 27 and 35, 28 and 36). Follow-up results from both time points are reported in the table. Five studies were excluded because of shorter follow-up period than 1 year.38-42

The follow-up period varied between 1 and 15 years. In all but two studies,31,37 there were no incentives for training in the follow-up period. Kiss et al.31 reported that the participants were told to continue training, and that reminders were used to incentive PFMT during the follow-up period. Kim et al.37 provided monthly group training classes, and asked the women to do individual home training. In most studies, loss to follow-up was reported, and varied between 0%2,27 and 39%.15 Adherence reported as number of women doing PFMT varied between 10%26 and 70%.27 Six of 17 studies did not report adherence to PFMT at follow-up or during the follow-up period.11,22,28,31,34,36

Most of the studies used self-report questionnaires for outcome assessment. Eight long-term studies22,23,27,28,33,34,36 interviewed the patients and/or used different pad tests, tested PFM function or applied urodynamic assessments. Eight of the studies used instruments that have been tested for reliability and validity, for example, ICIQ, Leakage index, Severity index, 7 day bladder diary.27,29,31-35,37 Twelve long-term follow-up studies reported surgery rates occurring in the follow-up period.20-21,26-28,34,35

Long-term results are shown in Table 1. Because of high heterogeneity in study design, outcome measures, cross-over of interventions, length of follow-up and losses to follow-up, no meta-analysis was performed. The results at long-term vary between studies. Surgery rates at follow-up vary between 4.9% at 28 months28 and 58% after 4-8 years.23 In the two studies with the longest follow-up, surgery rates were 8% at 10 years50 and 50% at 15 year.35 Only one RCT originally compared PFMT with surgery.23 After the initial intervention, which showed that surgery was superior to PFMT, the women were offered the other intervention. At follow-up, the initial satisfaction and cure rates were maintained in both the PFMT and surgery group. Bø et al.35 found that operated women were more likely to report severe incontinence (P = 0.03) and leakage that interfered with daily life (P = 0.04) than non-operated women at 15 year follow-up.

Altogether five studies stated that the initial success rate was maintained at follow-up23,24,29,32,33 Seven studies reported long-term outcome based on short-term success.22,23,28,30,34-36 All of these studies reported that the effect was better maintained in the responders than non-responders to the original program, and long-term success after short-term success varied between 41% and 85%. Kondo et al.28 reported that 19% of non-responders to short-term training were successors at 28 months follow-up, not counting the 4.9% who had surgery. In a later 8 years, follow-up by the same research group, the increase in muscle strength during the original program was the only reported parameter predicting positive long-term effect.30 No side effects from long-term PFMT have been reported.

DISCUSSION

This systematic review found 19 long-term studies on PFMT for women with SUI or MUI with predominately SUI symptoms. However, it is difficult to make meaningful comparisons between studies and to give pooled long-term cure rates, as the original short-term studies are heterogeneous when it comes to inclusion criteria, research design, outcome measures, exercise protocols with a huge variety of training dosages, use of adjuncts to PFMT such as biofeedback or vaginal cones, different adherence rates and finally different short-term success rates. For the long-term studies, further heterogeneity is added on in terms of length of the follow-up period, use of different outcome measures, co-interventions during the follow-up, competing events and losses to follow-up. This introduces what we would name “a double heterogeneity problem” in critical appraisal of long-term follow-up studies.

As for now, there are several recommendations on how to assess methodological quality of single RCTs19,43 and systematic reviews and meta-analysis,44 but we have not been able to find any specific guidelines on quality assessment of long-
<table>
<thead>
<tr>
<th>Author/year</th>
<th>Original design/Numbers (n)</th>
<th>Type of UI/PEDro score PEDro</th>
<th>Original intervention</th>
<th>Long-term follow-up period</th>
<th>Follow-up intervention?</th>
<th>Long-term outcome measure</th>
<th>Adherence not reported</th>
<th>Long-term effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferguson et al. (1990)</td>
<td>RCT/n = 20 SUI based on history and urodynamic assessment PEDro: 5/10</td>
<td></td>
<td>Group 1 (n = 10): 6 weeks of PFMT at home with a vaginal balloon as biofeedback; Group 2 (n = 10): As group 1 without biofeedback Vaginal palp</td>
<td>1 year</td>
<td>No</td>
<td>Questionnaire by letter or phone: Improvement of SUI and surgery rate</td>
<td>1 loss to follow-up: 5% reported to exercise</td>
<td>S had surgery (15%); None out of 29 were worse; Those still exercising reported to be improved</td>
</tr>
<tr>
<td>Cammu et al. (1991)</td>
<td>Cohort/n = 52 SUI based on history and urodynamic assessment</td>
<td></td>
<td>10 weeks of PFMT 30 min with PF twice/week - training as frequent as possible at home and use pre-contract. with cough, Diary Vaginal palp</td>
<td>14 months</td>
<td>No</td>
<td>Postal questionnaire with multiple choice and open questions: Improvement of SUI</td>
<td>7 losses to follow-up: Adherence not reported</td>
<td>Cured: 20% (9 of 45); Much improved: 38% (17 of 45); Some: 31% (14 of 45); Unchanged/worse: 11% (5 of 45); 3 had surgery (5.8%)</td>
</tr>
<tr>
<td>Mouritsen et al. (1991)</td>
<td>Cohort/N = 76; Women referred to surgery for SUI based on history and urodynamic assessment</td>
<td></td>
<td>3 months with PF. Individual instruction, self palpation 45 min weekly group sessions. Anal pressure measurement</td>
<td>1 year</td>
<td>No</td>
<td>Clinical assessment (?) with self report: Improvement of SUI = Pad test (unclear if this is included in follow-up cure rate)</td>
<td>Not reported</td>
<td>Cured: 30%; Much improved: 17%; 47% avoided surgery. 11 of 13 (85%) originally cured still cured. Of 37 patients improved 1/3 moved to cured</td>
</tr>
<tr>
<td>Klarickov et al. (1991)</td>
<td>RCT with cross over to surgery or PFMT after short-term period/ N = 52; SUI based on urodynamic assessment, PEDro: 6/10</td>
<td></td>
<td>Group 1 (n = 24): PFMT for 4 months with weekly group session with PF = home exercise in 4 positions with 5 contractions 4 times/day; Group 2 (n = 28): Surgery. Procedure chosen on basis of colposcopy-urethrography</td>
<td>Median 6 years (4–8)</td>
<td>14 out of 24 (58.3%) in the original PFMT group had surgery. 8 out of 28 (28.6%) in the originally operated had PFMT</td>
<td>Clinical assessment, interview or questionnaire on use of pads, improvement of SUI, number of incontinent patients, pad tests in 41, urinary diary in 37</td>
<td>4 loss to follow-up: 5% PFMT ≥ once a week, 28% occasionally, 14% never</td>
<td>10 had PFMT only. 2 of them worse, the rest were similar as short term. Some number of pads and incontinence episodes; 20 had surgery only 22 had both treatments; Same satisfaction rates as after original intervention period</td>
</tr>
<tr>
<td>Dougherty et al. (1993)</td>
<td>Cohort/n = 80 PFMT started after control period. Mild to moderate SUI based on history and urodynamic assessment</td>
<td></td>
<td>Four week control period (no treatment); 16 weeks of PFMT 3 times/week with nurse Palp. Measurement of PFM strength</td>
<td>14–26 months</td>
<td>No</td>
<td>Postal questionnaire asking about degree of urine loss</td>
<td>10 (15%) loss to follow-up at long term, 54% continued PFMT, 17% exercised 3 ≥/week</td>
<td>87% reported same urine loss as after cessation of training in original study or that it had diminished</td>
</tr>
<tr>
<td>Hahn et al. (1993)</td>
<td>Cohort with control group/n = 197 women referred to surgery for SUI based on history/urodynamic assessment Comparison group on waiting list for surgery PEDro: 3/10</td>
<td></td>
<td>Individual training program with PF once a week for 5 weeks, vaginal palp, then exercise at home 6–times/day for mean 4.7 months (range 1–18): Control group on waiting list for surgery</td>
<td>23% cured; 48% improved; 29% unchanged, 64% cured/ improved on provocative test (coughing, jumping, running)</td>
<td>2–7 years</td>
<td>No</td>
<td>Postal questionnaire asking about improvement of SUI</td>
<td>11% loss to follow-up: 15% daily training</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Author/year</th>
<th>Original design/Numbers (n)</th>
<th>Type of UI</th>
<th>PEDro score</th>
<th>Original intervention</th>
<th>Short-term effect</th>
<th>Long-term follow-up period</th>
<th>Follow-up intervention?</th>
<th>Long-term outcome measure</th>
<th>Loss to follow-up/adherence of PFMT in follow-up period</th>
<th>Long-term effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holley et al. (1995)</td>
<td>Cohort/n = 144 after Tchou et al. (1988); SUI, based on urodynamic assessment</td>
<td></td>
<td></td>
<td>4 weeks with individual PFMT sessions 30 min./2 times a week with PT</td>
<td>All subjects had subjective improvement. 9 (64.3%) had no leakage on cough or strain test</td>
<td>5 years</td>
<td>No</td>
<td>Postal questionnaire</td>
<td>4 (14.5%) loss to follow-up</td>
<td>4/10 had surgery and were continent 6/10 without surgery were leaking</td>
</tr>
<tr>
<td>Be and Talseth (1996)</td>
<td>RCT/n = 52 SUI after urodynamic assessment, PEDro 6/10</td>
<td></td>
<td></td>
<td>Group 1 (n = 30): Home exercise: vaginal palp, measurement of PFM strength, 7 visits with PT, motivation and measurement of PFM strength, 8–12 contractions 3 times/day Diary. Group 2 (n = 23): Intensive exercise: Same as home exercise + group training once a week. Emphasis on maximum contraction</td>
<td>Home exercise: 37% satisfied; Intensive exercise: 60% satisfied and positive closure pressure during cough. Significantly more reduction of leakage on pad test, social activity index and leakage index than home exercise.</td>
<td>5 years</td>
<td>No</td>
<td>Clinical assessment, urodynamics, pad testing, interview, leakage index, social activity index, measurement of PFM strength</td>
<td>Only intensive training group reported. No loss to follow-up. 70% satisfied, did not want further treatment, same muscle strength, 15 (75%) no visible leakage during cough, 11 had positive UCP during cough, 30% no leakage on pad test. Significant increase in leakage on pad testing and leakage index, but no change in social activity index. 3 (13.6%) had surgery, two successful, one had 17 g leakage on pad test after surgery</td>
<td>Follow-up of Intensive group only (n = 23). 70% satisfied, did not want further treatment, same muscle strength, 15 (75%) no visible leakage during cough, 11 had positive UCP during cough, 30% no leakage on pad test. Significant increase in leakage on pad testing and leakage index, but no change in social activity index. 3 (13.6%) had surgery, two successful, one had 17 g leakage on pad test after surgery</td>
</tr>
<tr>
<td>Glavind et al. (1996)</td>
<td>RCT/n = 40; SUI based on urodynamic assessment, PEDro 6/10</td>
<td></td>
<td></td>
<td>Vaginal palp Group 1: 2–3 times with individual instruction. Home training at least 3 times/day. Group 2: same as group one + 4 times with biofeedback (EMG) Sign better results subjectively and on pad test for group two. 58% cured in group two, 20% in group one</td>
<td>2–3 years</td>
<td>No</td>
<td>Postal questionnaire, improvement of SUI</td>
<td>3% loss to follow-up</td>
<td>Group two: 26% reported to be cured, 42% improved, 75% accept current situation, Group one: 0 reported to be cured, 29% improved, 50% accept current situation</td>
<td></td>
</tr>
<tr>
<td>Kondo and Yamada (1996)</td>
<td>Cohort/N = 103 SUI and MUI based on history. Study group divided on age over 65 years (n = 15) and under 65 years (n = 108)</td>
<td></td>
<td></td>
<td>Vaginal palpation. Rest and long contractions 90 min group training once a week with 10 participants for 8 weeks. Use of different postures in training group. Home exercise: 3 sets of 10 slow and 5 fast contractions + vaginal cones 15 min twice daily at home. Diary. Success rate (cured or reduction of UI &gt; 50%) 40% in group under 65 years and 20% in group over 65. Urine loss and bothersome scores improved in the youngest group only (pad test: from 15.00 g SD 14.9) to 5.9 g (SD 10.1). 95% stated that training was valuable and would recommend it</td>
<td>Mean 28 months (range 12–52)</td>
<td>No</td>
<td>Clinical assessment: Pad test, bother, self-reported success rate, muscle strength by manometer</td>
<td>No loss to follow-up. Adherence not reported</td>
<td>Same success rate 6 (4.9%) had surgery</td>
<td></td>
</tr>
<tr>
<td>Lagro-Janssen and van Weel (1998)</td>
<td>Cohort with control group n = 120, age 20–65 years, SUI, UUI, MUI based on history/</td>
<td></td>
<td></td>
<td>SUI: PFMT by GP, vaginal palp, written instruction, 5–10 daily sessions of 10 exercises; 60% dry/mildly incontinent in PFMT, one in control; 74% improved or cured, leakage episodes: PFMT: 20–7;</td>
<td>5 years</td>
<td>No</td>
<td>Postal questionnaire with self-report + 7 days bladder chart</td>
<td>5 loss to follow-up + 7 who had surgery; SUI adherence: 39% exercised ≥1/day, 15%</td>
<td>7 had surgery (6.9%), 14 had received additional treatment. Number of continent women</td>
<td></td>
</tr>
</tbody>
</table>

[Continued]
<table>
<thead>
<tr>
<th>Author/year</th>
<th>Original design/Numbers (n type of UI)</th>
<th>PEDro score</th>
<th>Short-term effect</th>
<th>Long-term follow-up period</th>
<th>Follow-up intervention?</th>
<th>Long-term outcome measure</th>
<th>Loss to follow-up/adherence of PFMT in follow-up period</th>
<th>Long-term effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cammu et al. (2000)³⁹</td>
<td>Cohort/n = 52; SUI based on history and urodynamic assessment</td>
<td>See Cammu et al. (1991)</td>
<td>See Cammu et al. (1991)</td>
<td>10 years</td>
<td>No</td>
<td>Postal questionnaire, review of medical files for surgery rates, self-reported improvement of SUI</td>
<td>1% loss to follow-up; 76% of those successful at short time had exercised, 55% of those not successful</td>
<td>16/24 (66.7%) successful patients remained satisfied; 2 (8%) had surgery</td>
</tr>
<tr>
<td>Kiss et al. (2002)³⁴ (abstract)</td>
<td>Cohort/n = 36; SUI based on history and urodynamic assessment</td>
<td>6 weeks PFMT by PT, assisted by biofeedback and electrical stimulation. Home training with PT controls every 3 months</td>
<td>Pad test: 10 dry; 17 improved; 2 unchanged; 2 worse; PFMT strength × time: 171.5 cm H₂O sec</td>
<td>6 years</td>
<td>Controls by PT and physician every 3 months</td>
<td>Clinical assessment Pad test, UPP, PFM strength (vaginal manometry), QoL (results not reported)</td>
<td>39% loss to follow-up (4 surgery, 8 lack of adherence, 2 dead)</td>
<td>Pad test: 6 dry; 12 improved; 4 unchanged; Strength × time: 120.6 cm H₂O sec</td>
</tr>
<tr>
<td>Alewijnse et al. (2003)³²</td>
<td>KCT/n = 121; SUI/urge/mixed based on history, PEDro 4/10</td>
<td>Four groups Group 1 (n = 29): Individual PFMT, Group 2–4 (n = 22, 25, 27): Individual PFMT + 1 out of 3 different adherence strategies: Diary</td>
<td>No difference between groups</td>
<td>1 year</td>
<td>No</td>
<td>Postal questionnaire including 7 day diary, improvement of SUI, wet episodes</td>
<td>20% loss to follow-up; 67% followed behavioral advice</td>
<td>Same results as short term</td>
</tr>
<tr>
<td>Parkkinen et al. (2004)³⁵</td>
<td>Cohort with two different groups/n = 40; SUI based on history and urodynamic assessment</td>
<td>Group 1 (n = 16): Outpatient training once per week for up to 1 year with PT including electrical stimulation and biofeedback + home exercise with 5 high intensity contractions for 5–8 sec + 5 low intensity sustained contractions for 20–30 sec twice a day, 5 days a week for at least 4 months. Group 2 (n = 17): Home training as above including cones but not electrical stimulation</td>
<td>No difference between groups</td>
<td>5 years</td>
<td>No</td>
<td>Clinical examination: pad test and questionnaire including urinary incontinence severity score (USS), self-reported cure/ improvement of SUI Assessment of PFM function with EMG</td>
<td>4 had surgery + one lost to follow-up; 25% exercised regularly after 12 months; 41% had physiotherapy in group 2 after 12 months</td>
<td>Group 1: 37.5% reported cure; 31.3% improvement; Group 2: 11.8% cure, 47.1% improved; Pad test reduction from mean 23 g to 14 g to 1 g in both groups; Significant improvements in strength in both groups. In general effect maintained</td>
</tr>
</tbody>
</table>

(Continued)
| Author/year | Original design/Numbers | Original intervention | Short-term effect | Long-term effect | Follow-up intervention? | Long-term outcome measure | Loss to follow-up/adherence of PFMT in follow-up period | Long-term effect | Studies are listed in chronological order. | MUI, mixed urinary incontinence. Palp, vaginal palpation. PMT, pelvic floor muscle training. PT, physical therapy. SUI, stress urinary incontinence. UPP, urethral pressure profile. UUI, urgency urinary incontinence. PEDro methodological quality score: random allocation, concealed allocation, baseline comparability, blinded assessors, blinded therapists, adequate follow-up (≥85%), results reported as point estimates and variability to treat. Between-group comparisons, results reported as point estimates and variability. | 
|-------------|--------------------------|-----------------------|------------------|-------------------|------------------------|--------------------------|-------------------------------------------------|----------------|--------------------------------------|
| Aukee et al. (2004) | RCT/n = 34; SUI based on history and urodynamic assessment; PEDro: 5/10 | All participants had 5 visits with PT during 12 weeks. All registered adherence. In addition: Group 1 (n = 19): home exercise 20 min 5 times/week; Group 2 (n = 16): individual EMG-assisted biofeedback device for home training | No difference between groups in Leakage index or pad test. Group 2 had better improvement in EMG activity | No difference between original groups in satisfaction or severity. 50% of each group had had surgery. Operated women more likely to report reduced bladder control with daily life of those satisfied 15 years ago | No | Postal questionnaire: number of pads used, ICIQ-UI SF, Severity index, leakage index | 4 (11.4%), but two were interviewed by phone. No report of adherence during follow-up period | 14% success rate. 19% of 46 successors were successes at 8 years. | 15 years | Advised to continue training on their own initiative |
| Bø et al. (2005) | RCT/n = 35; SUI based on history and urodynamic assessment; PEDro: 7/10 | Comparison of two different training regimens. See Bø and Talseth (1996) | No | Postal questionnaire: number of pads used, ICIQ-UI SF, Severity index, leakage index | 0% | No | No | Intensive exercise: 2 loss to follow-up; Home exercise: 4 loss to follow-up; No difference between groups in exercise adherence: 28% exercised at least once/week, 36% periodically, 36% never | 15 years | No |
| Kondo et al. (2007) | Cohort/n = 123; SUI and MUI based on history and urodynamic assessment; PEDro: 7/10 | 8 weeks PMT vaginal palp: nurses and physicians (60–90 min group training) + ‘perineal lock’, 30 repetitions/day at home, diary, weekly checks | 37% success rate at 6 weeks and 40% at 2 years | Clinical assessment: cough stress test, 60 min pad test, manometry. Postal questionnaire on self-reported cure/improvement | Mean 8 years (6–10) | No | Loss to follow-up: 36.8% 39% success rate. 19 of 46 successors were successors at 8 years (41%). 12 of 77 failures were successors at long-term (11.8%). The higher the improvement in muscle strength during initial training the better the results | Loss to follow-up 9.4% | No |
| Kim et al. (2007) | RCT with cross-over, hence no follow-up of control/n = 70 SUI based on ICIQ; PEDro: 6/10 | 60 min twice a week for 3 months, no vaginal palpation. PMT + general fitness class | Cured (leakage episodes): C3 54.5% of exercise group; C3 9.4% control; P < 0.001 | Interview based on ICIQ (leakage episodes): Yes | 1 year | YES; 1. group training class once a month. 2. Individual home training twice week 30 min PFM + other exercises | Loss to follow-up 37.1%; Adherence: Everyday: 30.3%; 2–3 times/week: 45.5%; Once/less week: 24.2% | 30.8% cure rate | 1 year | No |

**Legend:**
- **RCT:** Randomized Controlled Trial
- **SUI:** Stress urinary incontinence
- **MUI:** Mixed urinary incontinence
- **PMT:** Pelvic floor muscle training
- **PT:** Physical therapy
- **SUI:** Stress urinary incontinence
- **UPP:** Urethral pressure profile
- **UUI:** Urgency urinary incontinence
- **PEDro:** Physiotherapy Evidence Database
- **CI:** Confidence interval
- **SD:** Standard deviation
- **CI:** Confidence interval
- **SEM:** Standard error of the mean
- **CI:** Confidence interval
- **SEM:** Standard error of the mean

**References:**
term follow-up studies. Independent raters from the PEDro data-base had provided scores of methodological quality of the nine original short-term RCTs presented in this systematic review. As it is impossible to blind subjects and therapists during PFMT, eight should be considered the top-score for exercise studies. Scores between 4 and 6 can be considered moderate, and thus make a meaningful meta-analysis. However, this systematic review found that only one of the original RCTs compared PFMT with an untreated control, and that only five RCTs reported long-term effect according to the original treatment arms. These five trials were too heterogeneous to make a meaningful meta-analysis. In general, one may say that in spite of the fact that only two studies gave specific advice to continue PFMT or provided exercise classes during follow-up, some of the studies of PFMT showed surprisingly good long-term results assessed by self-report or surgery rates.

Eight of the studies had interviewed the patients and/or also conducted different clinical tests such as measurement of PFM function, pad testing or urodynamic assessments. Most of the studies used simple questionnaires and questions on satisfaction or improvement, but there were also use of instruments that had been tested for clinometric properties. Again, few studies had used the same outcome measures and if two studies had used the same, they were heterogeneous in other aspects, for example, design and interventions thus preventing meaningful comparison. As for surgery and drug studies, a combination of cure and improvement is often reported instead of absolute cure. Moreover, to date there is no consensus on what outcome measure to choose as the gold standard for cure (negative urethral closure pressure, number of leakage episodes, ≥2 g of leakage on pad test [tests with standardized bladder volume, 1, 24, and 48 hr], women’s report etc). In general, we would recommend that the same outcome measures should be used at both short- and long-term, and that only outcome measures that have been tested and found to be responsive, reliable and valid should be used in future follow-up studies.

As PFMT for SUI is considered a treatment to delay or avoid surgery, surgery rate in the follow-up period was chosen as our primary outcome measure of non-success. Surgery rates varied between 4.9% after 28 months and 58.3% after 4–8 years. Only one original RCT was found comparing the effect of surgery with PFMT, and short-term effect was clearly in favor of surgery. However, the short-term effect of both PFMT and surgery was maintained after 4–8 years. In the longest follow-up study, 50% in both originally randomized groups had had interval surgery. At 15 year follow-up, the short-term significant effect of the more intensive training protocol was no longer present. However, more women in the less intensive training group had surgery within the first 5 years after ending the training program. Interestingly, there were no differences in reported frequency or amount of leakage between non-operated or operated women, and women who had surgery reported significantly more severe leakage and to be more bothered by UI during daily activities than those not operated. There is, however, a selection bias to surgery, and the politics of when to offer surgery and to which women, vary widely between hospitals and countries. In addition, many women would not opt for surgery although they are incontinent. Hence, opting for surgery is a very difficult outcome measure to analyze and compare between studies. Hilton and Robinson have shown how cure rates of surgery vary widely with definitions and methods of measuring cure. For one surgical procedure cure rates varied between 9% and 85% depending on the definition of cure. We suggest that future long-term studies should involve both assessment of the actual leakage (pad tests and 3 day report of leakage episodes) and assessment of perceived impact and quality of life.

Obviously, long-term effect will depend on the initial success rate of an intervention as one would not expect short-term non-responders to be long-term responders. Hence, responders to the original trial might be the ones that should be in focus for long-term studies. This review found that only seven studies reported long-term outcome based on short-term success or non-success. All of these studies reported that the effect was better maintained among the responders than non-responders to the original program.

A common problem with follow-up studies after RCTs on PFMT is that usually women in the non-treatment or less effective intervention groups have received other interventions after cessation of the study period (cross-over or follow-up treatments). This may be supervised PFMT if they have been in the control group or medication or surgery if the patients wanted further treatment. If long-term results are reported following the original randomization and cross-over to other treatments is not taken into account, many women in the control group may have trained the PFM and comparison is no longer between training versus no training. Since many women may have cross-over or follow-up treatments, an intention to treat analysis at long term would bear little meaning. Further, there might be a power problem if analyzing only those who neither crossed-over nor had any follow-up treatments.

However, the main question is: can long-term outcome be expected after cessation of the active PFMT intervention? The effect of any training program will diminish with time if not continued or the pre- or co-contraction of the PFM has not reached an automatic level. In general, strength gain declines in a slower rate than the rate in which strength increases, but a 5–10% loss of muscle strength per week has been shown after training cessation. Greater losses have been shown in elderly (65–75 year olds) compared to younger (20–30 years old), and for both groups the majority of strength loss was from weeks 12 to 31 after cessation of training. The rate of strength loss may depend on length of the training period prior to detraining, type of strength test used and the specific muscle groups examined. Research has not yet indicated the exact resistance, volume, and frequency of strength training or the type of program needed to maintain training gains. However, studies indicate that to maintain strength gains or slow strength loss, the intensity should be maintained, but the volume and frequency of training can be reduced. One or 2 days a week seem to be an effective maintenance frequency for those individuals already engaged in a resistance training program.

So far, no studies have evaluated how many contractions subjects need to perform to maintain PFM strength after cessation of organized training. Lagro-Janssen and van Weel found that satisfaction was closely related to type of incontinence and adherence to training. Mixed incontinent women were more likely to lose the effect, and SUI women had the best long-term effect, but only 39% of them were exercising daily or "when needed." In some studies, the long-term effect seemed to be attributed to use of conscious pre-contraction before coughing and increase in intra-abdominal pressure.

To date, little is known about the long-term motivation for PFMT. Some women may find the exercises hard to conduct at a regular basis. However, Alewijnse found that most women followed advice of training 4–6 times a week 1 year after cessation of the training program. The following factors predicted adherence with 50%: positive intention to adhere, high...
short-term adherence levels, positive self-efficacy expectations, and frequent weekly episodes of leakage before and after initial therapy. In general, patients with different diseases do not comply with treatment for a wide variety of reasons: long lasting and time-consuming treatments, requirement of life-style changes, poor client/patient interaction, cultural and health beliefs, poor social support, inconvenience, lack of time, motivational problems, and travel time to clinics have been listed as factors for non-adherence.  

Strengths of the present systematic review are the comprehensive review of the literature based on both updated computerized search and use of published systematic reviews on short-term effect of PFMT.  

Due to published high quality systematic reviews of short-term effect studies in this area, we consider the risk of publication bias to be low. Limitations were the quality of individual studies, only one RCT comparing PFMT with no treatment, few reports of long-term effect following the original comparison groups, heterogeneity of interventions and outcome measures used, loss to follow-up, lack of reporting of co-interventions and cross-over and lack of reports of adherence, and incentives to follow-up training. These limitations will, however, also be present in long-term follow-up studies of surgery and medication interventions.  

There is a need for further high quality RCTs to evaluate the effect of different long-term incentives to continue PFMT after successful interventions. A possible way to maintain PFM strength after a treatment period is to include PFMT in general fitness classes for women. However, this will only involve those highly motivated for general fitness activities, and to date there is no knowledge about the effect of PFM maintenance training in fitness centers.

CONCLUSION

Nineteen long-term studies after PFMT were found. Meta-analysis of results was not possible due to high heterogeneity of both original and long-term studies. Long-term success based on responders to the original trial varied between 41% and 85%. Surgery rates at long term varied between 4.9% and 58%. Future high quality RCTs comparing different training dosages and follow-up strategies after cessation of short-term studies are warranted.

REFERENCES


