ABSTRACT

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 ≥ one year. Exclusion criteria were studies using electrical stimulation alone and studies in the peripartum period.

Results: Nineteen studies were included (1141 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.

Key words: exercise, follow-up, pelvic floor, urinary incontinence
INTRODUCTION

In 1948 Kegel (1) 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 the 1980s was there 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 (2,3,4,5,6,7).

Subjective cure/improvement rates of PFMT reported in RCTs in studies including groups with SUI and MUI vary between 56-70% (3,4,5,6,7). Short-term (immediately after cessation of training) cure rates of 44-80 %, defined as ≤ 2 grams of leakage on different pad tests, have been found after PFMT (8,9,10,11,12,13,14,15,16). The highest cure rates at short-term were shown in single blind RCTs of high methodological and interventional quality (14,15,16). 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 (14,15,16). Since biofeedback and electrical stimulation have not been conclusively shown to give additional effect to PFMT in RCTs and systematic reviews (3,4,5,7), 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 (6). 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 (urinary incontinence OR stress urinary incontinence) AND (follow-up OR long-term) with the following limits activated: humans, 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 was carried out (2,3,4,5,6,7,17).

Long-term was defined as ≥ one 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). 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 (19). PEDro is a 10 point scale giving one 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 seventeen long-term studies of PFMT fulfilling the inclusion criteria. Two additional studies were found by hand search of reference lists. The 19 studies included 1141 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 one year (38-42).

Nine of the long-term studies were based on an original pre-post (non-controlled) study design (21,22,24,26,28,30,31,33,36) whereas 9 studies were follow-up studies of original RCTs (11,20,23,27,29,32,34,35,37). One follow-up study was based on a non-randomized design with a control group (25). Mean PEDro score for the 9 RCTs was 5.1 (range 4-6).

Eight of the original RCTs providing long term follow-up studies compared different methods or intensities of PFMT while one RCT (29) and one non-randomized study (25) compared PFMT with untreated control groups and one RCT compared PFMT with surgery (23). In the two trials with an untreated control group, the control group crossed over to PFMT after the short term study period, and analyses of long term results between the original treatment groups could not be carried out.

The follow-up period varied between one 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 (21,27) to 39% (31). 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 studies (22,23,27,28,31,33,34,36) 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 e.g. 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 (20-23,26-30,34,35).

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 months (28) and 58% after 4-8 years (23). In the two studies with the longest follow-up, surgery rates were 8% at 10 years (30) 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=.03) and leakage that interfered with daily life (p= .04) than non-operated women at 15 year follow-up.

Altogether five studies stated that the initial success rate was maintained at follow-up (23,24, 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 eight 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 (36). 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 RCTs (19,43) and systematic reviews and meta-analysis (18), but we have not been able to find any specific guidelines on quality assessment of long-term follow-up studies. Independent raters from the PEDro database 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 (29), and that only five RCTs (11,23,32,34,35) 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 (31,37), some of the studies of PFMT showed surprisingly good long-term results assessed by self-report or surgery rates.

Eight of the studies (22,23,27,28,31,33,34,36) 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 e.g. design and interventions thus preventing meaningful comparison. As for surgery (44) and drug studies (45), 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 grams of leakage on pad test (tests with standardized bladder volume, 1 hour, 24 hour, and 48 hour), women’s report etc) (46,47). 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 (28) and 58.3% after 4-8 years (23). Only one original RCT was found comparing the effect of surgery with PFMT, and short term effect was clearly in favor of surgery (23). However, the short term effect of both PFMT and surgery was maintained after 4-8 years. In the longest follow-up study (35), 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 (47) 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 (46,47).

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 (22,23,28,30,34,35,36). 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 analysing only those who neither crossed-over nor had any follow-up treatments (28).

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 (48). 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 week 12-31 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 (48). One or two days a week seem to be an effective maintenance frequency for those individuals already engaged in a resistance training program (48).
So far, no studies have evaluated how many contractions subjects need to perform to maintain PFM strength after cessation of organized training. Lagro-Janssen & van Weel (29) 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 (27,30).

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 (49) found that most women followed advice of training 4-6 times a week one 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 (50).

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 (2-7). 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 (44,45). 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 centres.

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.
Reference List


17. Herbison P, Mantle J, Dean N. Weighted vaginal cones for urinary incontinence. The Cochrane Library 2007; 4


