Cochrane Library

**Cochrane** Database of Systematic Reviews

# Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women (Review)

Dumoulin C, Cacciari LP, Hay-Smith EJC

Dumoulin C, Cacciari LP, Hay-Smith EJC.
Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women.
*Cochrane Database of Systematic Reviews* 2018, Issue 10. Art. No.: CD005654.
DOI: 10.1002/14651858.CD005654.pub4.

**www.cochranelibrary.com**




WILEY

# TABLE OF CONTENTS

[Intervention Review]

# Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women

Chantale Dumoulin[1], Licia P Cacciari[1], E Jean C Hay-Smith[2]

[1]School of Rehabilitation, Faculty of Medicine, University of Montreal, Montreal, Canada. [2]Rehabilitation Teaching and Research Unit, Department of Medicine, University of Otago, Wellington, New Zealand

**Contact address:** Chantale Dumoulin, School of Rehabilitation, Faculty of Medicine, University of Montreal, C.P.6128 Succ. Centre-ville, Montreal, QC, H3C 3J7, Canada. chantal.dumoulin@umontreal.ca.

**Editorial group:** Cochrane Incontinence Group
**Publication status and date:** New search for studies and content updated (no change to conclusions), published in Issue 10, 2018.

**Citation:** Dumoulin C, Cacciari LP, Hay-Smith EJC. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2018, Issue 10. Art. No.: CD005654. DOI: 10.1002/14651858.CD005654.pub4.



## ABSTRACT

### Background

Pelvic floor muscle training (PFMT) is the most commonly used physical therapy treatment for women with stress urinary incontinence (SUI). It is sometimes also recommended for mixed urinary incontinence (MUI) and, less commonly, urgency urinary incontinence (UUI).

This is an update of a Cochrane Review first published in 2001 and last updated in 2014.

### Objectives

To assess the effects of PFMT for women with urinary incontinence (UI) in comparison to no treatment, placebo or sham treatments, or other inactive control treatments; and summarise the findings of relevant economic evaluations.

### Search methods

We searched the Cochrane Incontinence Specialised Register (searched 12 February 2018), which contains trials identified from CENTRAL, MEDLINE, MEDLINE In-Process, MEDLINE Epub Ahead of Print, ClinicalTrials.gov, WHO ICTRP, handsearching of journals and conference proceedings, and the reference lists of relevant articles.

### Selection criteria

Randomised or quasi-randomised controlled trials in women with SUI, UUI or MUI (based on symptoms, signs or urodynamics). One arm of the trial included PFMT. Another arm was a no treatment, placebo, sham or other inactive control treatment arm.

### Data collection and analysis

At least two review authors independently assessed trials for eligibility and risk of bias. We extracted and cross-checked data. A third review author resolved disagreements. We processed data as described in the *Cochrane Handbook for Systematic Reviews of Interventions*. We subgrouped trials by diagnosis of UI. We undertook formal meta-analysis when appropriate.

### Main results

The review included 31 trials (10 of which were new for this update) involving 1817 women from 14 countries. Overall, trials were of small-to-moderate size, with follow-ups generally less than 12 months and many were at moderate risk of bias. There was considerable variation in the intervention's content and duration, study populations and outcome measures. There was only one study of women with MUI and only one study with UUI alone, with no data on cure, cure or improvement, or number of episodes of UI for these subgroups.

*Symptomatic cure of UI at the end of treatment:* compared with no treatment or inactive control treatments, women with SUI who were in the PFMT groups were eight times more likely to report cure (56% versus 6%; risk ratio (RR) 8.38, 95% confidence interval (CI) 3.68 to 19.07; 4 trials, 165 women; high-quality evidence). For women with any type of UI, PFMT groups were five times more likely to report cure (35% versus 6%; RR 5.34, 95% CI 2.78 to 10.26; 3 trials, 290 women; moderate-quality evidence).

*Symptomatic cure or improvement of UI at the end of treatment:* compared with no treatment or inactive control treatments, women with SUI who were in the PFMT groups were six times more likely to report cure or improvement (74% versus 11%; RR 6.33, 95% CI 3.88 to 10.33; 3 trials, 242 women; moderate-quality evidence). For women with any type of UI, PFMT groups were two times more likely to report cure or improvement than women in the control groups (67% versus 29%; RR 2.39, 95% CI 1.64 to 3.47; 2 trials, 166 women; moderate-quality evidence).

*UI-specific symptoms and quality of life (QoL) at the end of treatment:* compared with no treatment or inactive control treatments, women with SUI who were in the PFMT group were more likely to report significant improvement in UI symptoms (7 trials, 376 women; moderate-quality evidence), and to report significant improvement in UI QoL (6 trials, 348 women; low-quality evidence). For any type of UI, women in the PFMT group were more likely to report significant improvement in UI symptoms (1 trial, 121 women; moderate-quality evidence) and to report significant improvement in UI QoL (4 trials, 258 women; moderate-quality evidence). Finally, for women with mixed UI treated with PFMT, there was one small trial (12 women) reporting better QoL.

*Leakage episodes in 24 hours at the end of treatment*: PFMT reduced leakage episodes by one in women with SUI (mean difference (MD) 1.23 lower, 95% CI 1.78 lower to 0.68 lower; 7 trials, 432 women; moderate-quality evidence) and in women with all types of UI (MD 1.00 lower, 95% CI 1.37 lower to 0.64 lower; 4 trials, 349 women; moderate-quality evidence).

*Leakage on short clinic-based pad tests at the end of treatment:* women with SUI in the PFMT groups lost significantly less urine in short (up to one hour) pad tests. The comparison showed considerable heterogeneity but the findings still favoured PFMT when using a random-effects model (MD 9.71 g lower, 95% CI 18.92 lower to 0.50 lower; 4 trials, 185 women; moderate-quality evidence). For women with all types of UI, PFMT groups also reported less urine loss on short pad tests than controls (MD 3.72 g lower, 95% CI 5.46 lower to 1.98 lower; 2 trials, 146 women; moderate-quality evidence).

Women in the PFMT group were also more satisfied with treatment and their sexual outcomes were better. Adverse events were rare and, in the two trials that did report any, they were minor. The findings of the review were largely supported by the 'Summary of findings' tables, but most of the evidence was downgraded to moderate on methodological grounds. The exception was 'participant-perceived cure' in women with SUI, which was rated as high quality.

**Authors' conclusions**

Based on the data available, we can be confident that PFMT can cure or improve symptoms of SUI and all other types of UI. It may reduce the number of leakage episodes, the quantity of leakage on the short pad tests in the clinic and symptoms on UI-specific symptom questionnaires. The authors of the one economic evaluation identified for the Brief Economic Commentary reported that the cost-effectiveness of PFMT looks promising. The findings of the review suggest that PFMT could be included in first-line conservative management programmes for women with UI. The long-term effectiveness and cost-effectiveness of PFMT needs to be further researched.

## PLAIN LANGUAGE SUMMARY

### Pelvic floor muscle training for urinary incontinence in women

**Review question**

We wanted to find out if pelvic floor muscle training (PFMT) helps women with urinary incontinence problems. We did this by comparing the effects of this training with no treatment, or with any inactive treatment (for example, advice on management with pads). We also summarised findings on costs and cost-effectiveness.

We searched for clinical trials up to 12 February 2018.

**Why is this question important?**

Stress incontinence is leaking of urine which cannot be easily controlled (if at all) when performing a physical activity. Physical activities could include coughing, sneezing, sporting activities or suddenly changing position. Urgency incontinence happens with a sudden, strong need to urinate. This can often lead to not making it to the toilet in time to urinate, resulting in leakage. Mixed incontinence is where someone has both stress and urgency incontinence.

PFMT is a programme of exercise to improve pelvic floor muscle strength, endurance, power, relaxation or a combination of these. It is a widely used treatment for women with stress, urgency and mixed incontinence.

**How did we carry out the review of the evidence?**

The 31 included trials involved 1817 women from 14 countries. The studies included women with stress, urgency or mixed urinary incontinence. The women were allocated randomly to either receive or not receive PFMT and the effects were compared. We looked at whether the condition was 'cured,' or 'cured or improved.' We also looked at symptoms, the effect on quality of life (QoL) and the frequency and amount of urine lost.

**Study funding sources**

Eight studies were publicly funded. Three received grants from public and private sources. Two received grants from private sources, while two studies received no funding. Sixteen studies did not declare their funding sources.

**What we found**

The quality of the evidence we looked at was mostly moderate, which means we can have some confidence in the results.

Cure of urinary incontinence after PFMT: women with stress urinary incontinence in the PFMT group were, on average, eight times more likely to report being cured. Women with any type of urinary incontinence in the PFMT group were, on average, five times more likely to report being cured.

Cure or improvement of urinary incontinence after PFMT: women with stress urinary incontinence in PFMT groups were, on average, six times more likely to report they were cured or improved. Women with all type of urinary incontinence in the PFMT group were roughly twice as likely to report they were cured or improved.

Leakage episodes after PFMT: women with stress urinary incontinence and women with all types of urinary incontinence in the PFMT group had one fewer leakage episode in 24 hours. PFMT appeared to reduce leakage episodes in women with urgency urinary incontinence alone.

For women with stress and all types of urinary incontinence, their incontinence symptoms and QoL were improved in the PFMT groups. Women were more satisfied with the PFMT treatment, while those in the control groups were more likely to seek further treatment.

Negative side effects of performing PFMT were rare and, in the two trials that did report them, the side effects were minor.

The authors of the one economic evaluation identified for the Brief Economic Commentary reported that the cost-effectiveness of PFMT looks promising.

**Summary of findings for the main comparison. Pelvic floor muscle training compared to control for stress urinary incontinence in women**

**Pelvic floor muscle training compared to control for stress urinary incontinence in women**

**Patient or population:** women with SUI
**Setting:** community-dwelling women
**Intervention:** PFMT
**Comparison:** no treatment, or inactive control treatments

| Outcomes | Anticipated absolute effects* (95% CI) | | Relative effect (95% CI) | № of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| | Risk with no treatment, placebo or control | Risk with PFMT | | | | |
| **Participant-perceived cure after treatment**<br>Treatment duration: 3–6 months | 60 per 1000[a] | 505 per 1000 (222 to 1000) | **RR 8.38** (3.68 to 19.07) | 165 (4 RCTs) | ⊕⊕⊕⊕ **High**[b] | — |
| **Participant-perceived cure or improvement after treatment**<br>Treatment duration: 3–6 months | 114 per 1000[a] | 720 per 1000 (442 to 1000) | **RR 6.33** (3.88 to 10.33) | 242 (3 RCTs) | ⊕⊕⊕⊖ **Moderate**[c,d] | — |
| **Number of leakage episodes in 24 hours**<br>assessed with: bladder diary<br>Treatment duration: 8 weeks to 6 months | The mean number of leakage episodes in 24 hours ranged from 1.07 to 3.61 episodes | MD 1.23 episodes lower (1.78 lower to 0.68 lower) | — | 432 (7 RCTs) | ⊕⊕⊕⊖ **Moderate**[e,f,g,h,i] | — |
| **Short (up to 1 hour) pad** test measured as grams of urine | The mean | MD 9.71 g lower (18.92 lower to 0.5 lower) | — | 185 (4 RCTs) | ⊕⊕⊕⊖ **Moderate**[i,j,k] | — |

| | | | | | |
|---|---|---|---|---|---|
| Treatment duration: 6 weeks to 6 months | short (up to 1 hour) pad test measured as grams of urine ranged from 3.64 to 38.70 g | | | | |
| **GRADE A UI-specific symptom measures**<br><br>Treatment duration: 4–12 weeks | 3 different Grade A psychometrically robust symptom questionnaires were used by trialists including KHQ severity domain (3 trials; n = 65), ICIQ-UI Short Form (3 trials; n = 98) and UDI (1 trial; n = 17). Participants in the PFMT group reported significant improvement in UI symptoms. | — | (7 RCTs) | ⊕⊕⊕⊖<br>**Moderate**[l,m,n,o] | — |
| **GRADE A UI-specific QoL measures**<br><br>Treatment duration: 6 weeks to 6 months | 5 different Grade A psychometrically robust QoL questionnaires were used by trialists including KHQ impact domain (3 trials; n = 65), KHQ physical limitation domain (3 trials; n = 65); ICIQ-LUTSqol (1 trial; n = 60); IIQ (1 trial; n = 17); I-QOL (1 trial; n = 24). Participants in the PFMT group reported significant improvement in UI-specific QoL except for the KHQ impact after treatment; however with considerable heterogeneity ($I^2$ = 76%). | — | (6 RCTs) | ⊕⊕⊖⊖<br>**Low**[i,l] | — |

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

**CI:** confidence interval; **I-QOL:** Incontinence of Quality of Life questionnaire; **ICIQ-LUTSqol:** ICIQ Lower Urinary Tract Symptoms Quality of Life; **ICIQ-UI Short Form:** Incontinence Modular Questionnaire Urinary Incontinence Short Form; **IIQ:** Incontinence Impact Questionnaire; **KHQ:** King's Health Questionnaire; **MD:** mean difference; **PFMT:** pelvic floor muscle training; **QoL:** quality of life; **RCT:** randomised controlled trial; **RR:** risk ratio; **SUI:** stress urinary incontinence; **UDI:** Urinary Distress Inventory; **UI:** urinary incontinence.

**GRADE Working Group grades of evidence**
**High quality:** we are very confident that the true effect lies close to that of the estimate of the effect.
**Moderate quality:** we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
**Low quality:** our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
**Very low quality:** we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Assumed risk based on number of events.

[b]Large RR and confidence interval in two trials.
[c]Random sequence generation and allocation concealment at high risk in one trial (Lagro-Janssen 1991a).
[d]Blinding of outcome assessor unclear in one trial, for which the participants filled web-based questionnaires with no face-to-face interaction with the researcher group (Asklund 2017).
[e]Random sequence generation and allocation concealment at high risk in one trial (Lagro-Janssen 1991a).
[f]Allocation concealment and incomplete outcome data unclear in three trials (Burns 1993; Firra 2013; McLean 2013).
[g]Blinding of outcome assessment unclear in two trials (Asklund 2017; McLean 2013), and at high risk in one trial (Firra 2013).
[h]Baseline comparability at high risk in one trial, but not for this outcome and not for this subgroup (urinary frequency for the urge incontinent subgroup only, Firra 2013).
[i]Considerable heterogeneity ($I^2$ higher than 75%).
[j]Random sequence generation unclear, and blinding of outcome assessment at high risk in one trial (Pereira 2011).
[k]Allocation concealment and incomplete outcome data unclear in two trials (McLean 2013; Pereira 2011).
[l]Downgraded for being considered a self-reported measure.
[m]Random sequence generation, allocation concealment, incomplete data and blinding of outcome assessor unclear for one trial (Carneiro 2010).
[n]Allocation concealment and incomplete outcome data unclear in one trial (Pereira 2011).
[o]Unclear for bias except baseline comparability and selective reporting (Beuttenmuller 2010).

**Summary of findings 2. Pelvic floor muscle training compared to control for urinary incontinence (all types) in women**

**Pelvic floor muscle training compared to control for urinary incontinence (all types) in women**

**Patient or population:** women with UI (all types)
**Setting:** community-dwelling women
**Intervention:** PFMT
**Comparison:** no treatment, placebo or control

| Outcomes | Anticipated absolute effects* (95% CI) | | Relative effect (95% CI) | № of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| | Risk with no treatment, placebo or control | Risk with PFMT | | | | |
| **Participant-perceived cure after treatment**<br>Treatment duration: 8–12 weeks | 62 per 1000[a] | 329 per 1000<br>(171 to 632) | **RR 5.34**<br>(2.78 to 10.26) | 290<br>(3 RCTs)[a] | ⊕⊕⊕⊖<br>**Moderate**[b,c,d] | — |
| **Participant-perceived cure or improvement after treatment**<br>Treatment duration: 6–8 weeks | 288 per 1000[a] | 687 per 1000<br>(471 to 998) | RR 2.39<br>(1.64 to 3.47) | 166<br>(2 RCTs)[a] | ⊕⊕⊕⊖<br>**Moderate**[e,f,g] | — |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Number of leakage episodes in 24 hours** assessed with: bladder diary<br>Treatment duration: 8–12 weeks | The mean number of leakage episodes in 24 hours ranged from 1.06 to 2.50 | MD 1 episode lower (1.37 lower to 0.64 lower) | — | 349 (4 RCTs) | ⊕⊕⊕⊖ **Moderate**[h,i,l] | — |
| **Short (up to 1 hour) pad test measured as grams of urine**<br>Treatment duration: 6 weeks to 6 months | The mean short (up to 1 hour) pad test measured as grams of urine ranged from 5.10 g to 8.40 g | MD 3.72 g lower (5.46 lower to 1.98 lower) | — | 146 (2 RCTs) | ⊕⊕⊕⊖ **Moderate**[i,j] | — |
| **GRADE A UI-specific symptom measures**<br>Treatment duration: 12 weeks | 1 Grade A psychometrically robust symptom questionnaire was used by 1 trial (n = 63); the UDI. Participant in the PFMT group reported significant improvement in UI-specific symptoms. | | — | (1 RCT) | ⊕⊕⊕⊖ **Moderate**[g,i,k] | — |
| **GRADE A UI-specific QoL measures**<br>Treatment duration: 6–12 weeks | 4 different Grade A psychometrically robust QoL questionnaires were used by trialists including the IIQ short form (2 trials; n = 91), the IIQ long form (1 trial; n = 24); I-QOL (1 trial; n = 17). Participant in the PFMT group reported significant improvement in UI-specific QoL. | | — | (4 RCTs) | ⊕⊕⊖⊖ **Low**[g,i,l,m] | — |

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

**CI:** confidence interval; **I-QOL:** Incontinence of Quality of Life questionnaire; **IIQ:** Incontinence Impact Questionnaire; **KHQ:** King's Health Questionnaire; **MD:** mean difference; **PFMT:** pelvic floor muscle training; **QoL:** quality of life; **RCT:** randomised controlled trial; **RR:** risk ratio; **UDI:** Urinary Distress Inventory; **UI:** urinary incontinence.

**GRADE Working Group grades of evidence**
**High quality:** we are very confident that the true effect lies close to that of the estimate of the effect.
**Moderate quality:** we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
**Low quality:** our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
**Very low quality:** we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Assumed risk based on number of events.
[b]Allocation concealment unclear in two trials (Burgio 1998 which was the biggest trial, and in Kim 2007).
[c]Incomplete outcome data and blinding of outcome assessor unclear for two trials (Kim 2007; Kim 2011a).
[d]Considerable heterogeneity ($I^2$ higher than 75%).
[e]Allocation concealment unclear in both trials (Burgio 1998; Diokno 2010).
[f]Baseline comparability at high risk for one trial, with older participants in the PFMT group (Diokno 2010).
[g]Downgraded for being considered a self-reported measure.
[h]Allocation concealment unclear in one trial (Burgio 1998).

[i]Incomplete outcome data, blinding of participant and personnel, baseline comparability for a different outcome unclear in Celiker Tosun 2015 (PFMT group presenting lower impact on quality of life and higher night-time urinary frequency).
[j]Random sequence generation and allocation concealment unclear in one trial (Yoon 2003).
[k]Only one trial of a small sample size.
[l]Blinding of outcome assessor and baseline comparability at high risk in one trial (PFMT group older (P = 0.06) and presenting higher impact on quality of life (P = 0.06); Leong 2015).
[m]High risk for incomplete outcome data and blinding of outcome assessment (Sar 2009).

### Summary of findings 3. Pelvic floor muscle training compared to control for urgency urinary incontinence in women

**Pelvic floor muscle training compared to control for urgency urinary incontinence in women**

**Patient or population:** UUI in women
**Setting:** community-dwelling women
**Intervention:** PFMT
**Comparison:** no treatment, placebo or control

| Outcomes | Anticipated absolute effects* (95% CI) | | Relative effect (95% CI) | № of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| | Risk with no treatment, placebo or control | Risk with PFMT | | | | |
| **Participant-perceived cure after treatment** | — | — | — | (0 studies) | — | — |
| **Participant-perceived cure or improvement after treatment** | — | — | — | (0 studies) | — | — |
| **Number of leakage episodes in 24 hours** assessed with: bladder diary<br>Treatment duration: 8 weeks | The mean number of leakage episodes in 24 hours was 2.60 | MD 1.83 episodes lower (2.65 lower to 1.01 lower) | — | 12 (1 RCT) | ⊕⊕⊖⊖ **Low**[a,b] | — |
| **Short (up to 1 hour) pad test** measured as grams of urine | — | — | — | (0 studies) | — | — |
| **GRADE A UI-specific symptom measures** | — | | — | (0 studies) | — | — |
| **GRADE A UI-specific QoL measures** | — | | — | (0 studies) | — | — |

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

**CI:** confidence interval; **MD:** mean difference; **PFMT:** pelvic floor muscle training; **QoL:** quality of life; **RCT:** randomised controlled trial; **UI:** urinary incontinence; **UUI:** urgency urinary incontinence.

**GRADE Working Group grades of evidence**
**High quality:** we are very confident that the true effect lies close to that of the estimate of the effect.
**Moderate quality:** we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
**Low quality:** our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
**Very low quality:** we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

---

[a]Blinding of outcome assessor and baseline comparability for a different outcome at high risk (urinary frequency higher for the PFMT group, Firra 2013).
[b]Block randomisation not properly specified, and allocation concealment and attrition bias unclear (Firra 2013).

# BACKGROUND

## Description of the condition

### Urinary incontinence

Urinary incontinence (UI) is a common problem among adults living in the community. It is more frequent in women, increasing with age, and is particularly common among those in residential care (Hunskaar 2002). Estimates of prevalence are influenced by the definition of incontinence, the sample population and the format of questions about incontinence. In addition, figures are unlikely to reflect the true scope of the problem because embarrassment and other factors may lead to under-reporting. Estimates of the prevalence of UI in women vary between 25% to 45% in most studies (Milsom 2013). Data from the widely cited EPINCONT study of UI in women (27,936 Norwegian women) suggest a gradual increase in prevalence with age to an early peak prevalence around midlife (50 to 54 years), which coincides with menopause, followed by a slight decline or stabilisation until about 70 years of age, when the prevalence begins to rise steadily (Hannestad 2000). Pregnancy, labour and vaginal delivery (versus caesarean section) are significant risk factors for later UI, but the strength of this association diminishes substantially with age (Milsom 2013).

Isolated stress urinary incontinence (SUI) accounts for half of all UI, with most studies reporting 10% to 39% prevalence. With few exceptions, mixed urinary incontinence (MUI) is the next most common, with most studies reporting 7.5% to 25% prevalence. Isolated urgency urinary incontinence (UUI) is uncommon, with 1% to 7% prevalence (Milsom 2013). The type of urine leakage is classified according to what is reported by the woman (symptoms), what is observed by the clinician (signs) and on the basis of urodynamic studies. The definitions of the different types of UI given below are those of the International Continence Society (Haylen 2010).

Not only is UI a serious medical condition in that it can lead to perineal rash, pressure ulcers and urinary tract infections (Resnick 1989), it is also an undeniable social problem, creating embarrassment and negative self-perception (Papanicolaou 2005). UI reduces both social interactions and physical activities (Resnick 1989), and is also associated with poor self-rated health (Johnson 1998), impaired emotional and psychological well-being (Coyne 2012), and impaired sexual function (Sen 2006). In the medium- or long-term, women with UI often find themselves isolated and relatively inactive (Fantl 1996). Moreover, UI in older women doubles the risk of admission to a nursing home, independent of age or the presence of comorbid conditions (Hunskaar 1991). Current estimates of the costs of UI are not available. In 1995, the costs of UI in women was estimated at USD 12.4 billion in the US, suggesting that its economic burden is considerable (Wilson 2001).

### Stress urinary incontinence

If a woman reports involuntary urine leakage with physical exertion or a clinician observes urine leakage at the same time as the exertion, this is called SUI. When urodynamic studies demonstrate involuntary loss of urine during increased intra-abdominal pressure, but the leakage is not accompanied by a contraction of the detrusor muscle (bladder smooth muscle), this is called urodynamic SUI (Haylen 2010). SUI is likely to be due to anatomical defects in the structures that support the bladder and urethra, resulting in suboptimal positioning of these structures at rest or on exertion, or dysfunction of the neuromuscular components that help control the urethral sphincter or urethral pressure. As a result, the bladder outlet (urethra) is not closed off properly during exertion, which results in leakage.

### Urgency urinary incontinence

The symptom of UUI is present when a woman reports involuntary leakage associated with or immediately preceded by a sudden, compelling need to void (urgency). The sign of UUI is identified by the observation of involuntary urine leakage from the urethra synchronous with the sensation of a sudden, compelling desire to void that is difficult to defer. UUI usually results from an involuntary increase in bladder pressure due to the contraction of the detrusor muscle. If there is a known neurological cause for the detrusor muscle dysfunction, this is called neurogenic detrusor overactivity. If the cause is not known, the condition is called idiopathic detrusor overactivity.

### Mixed urinary incontinence

Many women have symptoms or signs of both SUI and UUI, and urodynamic studies sometimes reveal that urine leakage results from a combination of urodynamic SUI and detrusor overactivity. When women have either or both symptoms and signs of SUI and UUI, this is called MUI.

## Description of the intervention

### Treatment of urinary incontinence

A wide range of treatments have been used in the management of UI, including conservative interventions (such as physical therapies, lifestyle interventions, behavioural training and anti-incontinence devices), pharmaceutical interventions and surgery. This review will focus on one of the physical therapies, specifically pelvic floor muscle training (PFMT).

### Pelvic floor muscle training

PFMT has been part of exercise programmes in Chinese Taoism for over 6000 years (Chang 1984). It first entered modern medicine in 1936, when a paper describing tensing and relaxing of the pelvic floor muscles introduced the use of PFMT as a prevention and treatment option for urinary and faecal incontinence to the British physiotherapy profession (Morris 1936). However, PFMT as a treatment for SUI did not become widespread until after the mid-1900s when American gynaecologist Arnold Kegel reported on the successful treatment of 64 cases of female SUI using pelvic floor muscle exercises with a pressure biofeedback perineometer (Kegel 1948). More recently, PFMT was defined by an international standardisation committee as an exercise to improve pelvic floor muscle strength, endurance, power, relaxation or a combination of these parameters (Bø 2017).

## How the intervention might work

### Biological rationale for pelvic floor muscle training for stress and mixed urinary incontinence

The biological rationale is two-fold. First, an intentional, effective pelvic floor muscle contraction (lifting the pelvic floor muscles in a cranial and forward direction) prior to and during effort or exertion clamps the urethra and increases urethral pressure, preventing urine leakage (DeLancey 1988). Ultrasonography and magnetic resonance imaging (MRI) studies have demonstrated the cranial

and forward movement of the pelvic floor muscles during active contraction and the resulting impact on the urethral position, which supports this rationale (Bø 2001). Miller 1998 named this counter-balancing pelvic floor muscle contraction prior to a cough as the 'knack' and assessed its effectiveness in a randomised controlled trial (RCT). They demonstrated that a voluntary pelvic floor muscle contraction (VPFMC) before or during coughing can reduce leakage after only one week of training. Other published research, employing the term 'pelvic floor muscle functional training,' recommends precontracting the pelvic floor muscles not only during a cough but for any daily task that results in increased intra-abdominal pressure (Carrière 2006). Thus, research suggests that the timing of a pelvic floor muscle contraction might be an important factor in the maintenance of UI.

However, the optimal strength required to clamp the urethra and prevent urine leakage has not yet been determined. In healthy continent women, activation of the pelvic floor muscles before or during physical exertion seems to be an automatic response that does not require conscious effort (Peschers 2001). There is some evidence that this pelvic floor muscle 'reflex' contraction is a feed-forward loop and might precede a bladder pressure rise by 200 ms to 240 ms (Constantinou 1982). For women with UI, learning to rapidly perform a strong, well-timed pelvic floor muscle contraction may actively prevent urethral descent during an intra-abdominal rise in pressure (Bø 1995).

Second, the bladder neck receives support from strong, toned pelvic floor muscles (resistant to stretching), limiting its downward movement during effort and exertion, and thus preventing urine leakage (Bø 2004). Bø 2004 suggests that intensive strength training may build up the structural support of the pelvis by permanently elevating the levator plate to a higher position inside the pelvis and by enhancing the hypertrophy and stiffness of its connective tissues. In line with, and supporting this hypothesis, differences in the anatomical position of the pelvic floor muscles have been demonstrated between women with and without UI (Pontbriand-Drolet 2012). Additionally, dynamometric studies have shown that women with SUI or MUI demonstrate less pelvic floor muscle tone, maximal strength, rapidity of contraction and endurance as compared to continent women (Pontbriand-Drolet 2012).

Further, in an uncontrolled MRI reconstruction study, a significant reduction in the internal surface area of the levator ani was observed after PFMT, suggesting an increase in passive stiffness of the levator ani, which is indicative of the state of pelvic floor muscle tone (Dumoulin 2007). Griffin 1994, using a pressure probe inside the vagina, also showed a significant difference in women's pelvic floor muscle resting pressure three to four weeks after starting PFMT, and increased resting pressure after PFMT was completed. Furthermore, Balmforth 2004 reported increased urethral stability at rest and during effort following 14 weeks of supervised PFMT and behavioural modifications.

Thus, there is a growing body of evidence to support the rationale that PFMT improves pelvic floor muscle tone, which may facilitate more effective automatic motor unit firing of the PFM. This prevents pelvic floor muscle descent during increased intra-abdominal pressure, which in turn prevents urine leakage (Bø 2007). Given the above biological rationale, the objective of PFMT for SUI is usually to improve the timing (of contraction), strength, endurance and stiffness of the pelvic floor muscles.

### Biological rationale for pelvic floor muscle training for urgency urinary incontinence

PFMT can also be used in the management of UUI. The biological rationale is based on Godec's observation that a detrusor muscle contraction can be inhibited by a pelvic floor muscle contraction induced by electrical stimulation (Godec 1975). Further, de Groat 2001 demonstrated that during urine storage there is an increased pudendal nerve outflow response to the external urethral sphincter, increasing intraurethral pressure and representing what he termed a 'guarding reflex' for continence.

Additionally, Morrison 1995 demonstrated that Barrington's micturition centre excitatory loop switches on when bladder pressures are between 5 and 25 mmHg, while the inhibitory loop is predominantly active above 25 mmHg. Inhibition involves an automatic (unconscious) increase in tone for both the pelvic floor muscle and the urethral striated muscle. Thus, VPFMC may be used to control UUI. After inhibiting the urgency to void and the detrusor contraction, the woman can reach the toilet in time to avoid urine leakage. However, the number, duration, intensity and timing of the pelvic floor muscle contraction required to inhibit a detrusor muscle contraction is not known.

### Types of pelvic floor muscle training programmes

There is not an absolute dividing line that differentiates strength from endurance-type exercise programmes. It is common for both strength and fatigue resistance to improve in response to an exercise programme, although one may be affected more than another. Characteristic features of strength training include low numbers of repetitions with high loads, where ways to increase load include increasing the amount of voluntary effort with each contraction and performing exercise with and then against gravity. Endurance training is characterised by high numbers of repetitions or prolonged contractions with low-to-moderate loads. Behavioural training to improve co-ordination and urge suppression usually involves the repeated use of VPFMC in response to a specific situation, for example VPFMC prior to cough, and VPFMC with the sensation of urgency.

## Why it is important to do this review

Many women are referred for PFMT on the basis of symptoms or clinical signs of stress, urgency or mixed UI. There is currently no consensus about the need for urodynamic investigations before PFMT (Clement 2013), but a single RCT indicated that there was no evidence of a difference in the conservative treatment outcome if the referral was made on the basis of symptom diagnosis or urodynamics (Ramsay 1995). The sensitivity and specificity of urodynamic diagnosis seems variable depending on the expertise of the investigator, the scope of testing and the dysfunction being investigated. For these reasons, we included diagnoses based on symptoms, signs and urodynamic investigations in this review.

This is an update of Dumoulin 2014, which was first published in 2001 and last updated in 2014. This review investigates whether PFMT is an effective treatment in the management of female stress, urgency and mixed UI compared to no treatment, placebo, sham or control treatments. Women place high value on the resolution of symptoms associated with UI. Johannesson 1997 estimated the willingness to pay for a 50% reduction of symptoms in a group of women with UI. The mean willingness to pay for a 50% reduction in symptoms was GBP 89 per month (SEK 1030 in 1997). Due to

the chronic and long-standing nature of UI, knowing the costs and benefits of PFMT is important and will provide valuable information to decision makers.

Other reviews regarding UI in women and PFMT address whether:

- one type of PFMT is better than another (Hay-Smith 2011), or whether feedback or biofeedback has a role to play (Herderschee 2011);
- PFMT is better than other treatments (e.g. other physical therapies, medication and surgery) (protocol by Lins 2014); and
- if the addition of PFMT to other therapies adds benefit (Ayeleke 2015).

A separate review considers the role of PFMT in the treatment and prevention of urinary and faecal incontinence related to childbirth (Woodley 2017). An Overview of Cochrane systematic reviews is underway and aims to consider the role of conservative interventions for UI in women, one of which is PFMT (McClurg 2016).

Earlier Cochrane Reviews of PFMT (Dumoulin 2010; Dumoulin 2014; Hay-Smith 2002; Hay-Smith 2006), as well as other previously published systematic reviews of PFMT (Berghmans 1998; Berghmans 2000; Bø 1996; de Kruif 1996; Fedorkow 1993; Wilson 1999), are outdated as new trials have been published. Although these reviews identified several PFMT trials, there were few data and considerable clinical heterogeneity in the studies. There is sufficient uncertainty about the effects of PFMT, particularly the size of effect, to suggest that continuing to update earlier Cochrane Reviews is warranted.

## OBJECTIVES

To assess the effects of PFMT for women with UI in comparison to no treatment, placebo or sham treatments, or other inactive control treatments; and summarise the findings of relevant economic evaluations.

## METHODS

### Criteria for considering studies for this review

#### Types of studies

We included RCTs and quasi-randomised trials (e.g. using allocation by alternation). Other forms of controlled clinical trials were excluded.

#### Types of participants

We included trials of women with UI and diagnosed as having SUI, UUI or MUI on the basis of symptoms, signs or urodynamic evaluation, as defined by the trialists. Trials that recruited men and women were eligible for inclusion providing demographic and outcome data were reported separately for women.

We excluded trials of women with UI whose symptoms might be due to significant factors outside the urinary tract (e.g. neurological disorders, cognitive impairment, lack of independent mobility and cancer or radiotherapy). We excluded studies investigating nocturnal enuresis in women. We excluded studies that specifically recruited antenatal or postnatal women (childbearing women). Given the physiological changes of the pregnancy and postpartum period, it is possible that the effect of PFMT might differ in this group. PFMT for the prevention and management of UI in antenatal and postnatal women is addressed in another Cochrane Review (Woodley 2017).

#### Types of interventions

One arm of all eligible trials included a PFMT programme to ameliorate symptoms of existing urine leakage. Thus, we excluded studies including only asymptomatic women doing PFMT for primary or secondary prevention of UI. Another arm of the trial was a no treatment arm, a placebo treatment arm, a sham treatment arm (e.g. sham electrical stimulation) or an inactive control treatment arm (e.g. advice on the use of pads).

PFMT was defined as a programme of repeated VPFMC taught and supervised by a healthcare professional. We considered all types of PFMT programmes, including using variations in the purpose and timing of PFMT (e.g. PFMT for strengthening, and PFMT for urge suppression), different ways of teaching PFMT, types of contractions (fast or sustained), and number of contractions.

We included trials that combined PFMT with a single episode of biofeedback (for the purposes of teaching a pelvic floor muscle contraction) or advice on strategies for symptoms of urgency or frequency (but without a scheduled voiding regimen characteristic of bladder training). We excluded trials that combined PFMT with another conservative therapy (e.g. bladder training, biofeedback, vaginal cones or electrical stimulation) or drug therapy (e.g. an anticholinergic drug).

#### Types of outcome measures

A subcommittee (Outcome Research in Women) of the Standardisation Committee of the International Continence Society suggested that research investigating the effect of therapeutic interventions for women with UI consider five outcome categories: the woman's observations (symptoms), quantification of symptoms (e.g. urine loss), the clinician's observations (anatomical and functional), quality of life (QoL) and socioeconomic measures (Lose 1998). One or more outcomes of interest from each domain were chosen for this review.

The review authors also considered the International Classification of Function, Disability, and Health (ICF), a World Health Organization (WHO) initiative describing a conceptual framework for understanding health and the consequences of health conditions, when choosing the primary outcomes of interest for the review (WHO 2002). The framework describes the inter-relationships between a woman's impairment of body functions and structures (e.g. pelvic floor muscle dysfunction), limitations in activity (e.g. avoiding running because of leakage), and restricted participation (e.g. avoiding social activities with friends and family because of leakage). Thus, the choice of condition-specific QoL as one of the primary outcome measures reflects the importance the authors placed on the effects of incontinence on women's activities and participation, while a measure of impairment (e.g. of pelvic floor muscle function) was of secondary importance.

##### *Primary outcomes*

###### Participant-reported measures

- Symptomatic cure of UI at the end of treatment.
- Symptomatic cure or improvement of UI at the end of treatment.
- Symptom- and condition-specific QoL measures (e.g. King's Health Questionnaire (Kelleher 1997), Incontinence Quality of

Life (I-QOL) (Donovan 2005), Bristol Female Lower Urinary Tract Symptoms (B-FLUTS) questionnaire (Jackson 1996)).

### *Secondary outcomes*

#### Participant-reported measures

- Longer-term symptomatic cure and improvement (six months to one year after end of treatment; more than one year after end of treatment).
- Satisfaction.
- Need for further treatment (e.g. need of aids, surgery, drugs, PFMT).
- Self-efficacy (e.g. UI-specific self-efficacy scale (Tannenbaum 2009), or PFMT-specific self-efficacy scales (Broome 2003)).

#### Participant-reported quantification of symptoms

- Number of urinary leakage episodes (in 24 hours).
- Number of micturitions during the day (frequency).
- Number of micturitions during the night (nocturia).

#### Clinicians' measures

- Pad and paper towel testing short (up to one hour) or long (24 hours) urine loss (grams of urine lost) at the end of treatment.
- Number cured or improved based on pad weights in short clinic-based pad test at the end of treatment.
- Other pad or paper towel tests (e.g. those not reported as cure, cure and improvement or grams, those reported at other time points after treatment).

#### Quality of life (not condition specific)

- General health status measures (e.g. Short Form-36 (Ware 1993)).
- Psychosocial outcome measures (e.g. Hopkins Symptoms Checklist for psychological distress (SCL-90-R) (Derogatis 1974), Hospital Anxiety and Depression Scale (HADS) (Zigmond 1983)).
- Sexual function or problems (e.g. leakage during intercourse, impact on sexual function).

#### Adverse effects

- Adverse effects (e.g. discomfort, soreness, pain, bleeding).

#### Measures of likely moderator variables

#### Measures of pelvic floor muscle function

- Digital evaluation.
- Pelvic floor muscle dynamometry.
- Pelvic floor muscle electromyography.
- Vaginal squeeze pressure.
- Perineal ultrasound.

#### Measure of adherence

- Number of study participants attending or completing treatment sessions.
- Number of study participants performing PFMT or adherence to home and clinic-based PFMT.
- Number of contractions completed per session, day or week.

#### Main outcomes for 'Summary of findings' tables

In accordance with guidance in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), we selected the following outcomes for inclusion in the 'Summary of findings' tables:

- participant-reported symptomatic cure of UI at the end of treatment;
- participant-reported symptoms of cure or improvement of UI at the end of treatment;
- symptom and condition-specific QoL measures (Grade A UI-specific symptoms and QoL measures);
- number of urinary leakage episodes (in 24 hours); and
- short (up to one hour) pad test measured as grams of urine at the end of treatment.

## Search methods for identification of studies

This review drew on the search strategy developed by Cochrane Incontinence. There were no language or other restrictions on any of the searches described below.

### Electronic searches

We identified relevant trials from the Cochrane Incontinence Specialised Register. For more details of the search methods used to build the Specialised Register, please see the Group's webpages where details of the Register's development (from inception) and the most recent searches performed to populate the Register can be found. To summarise, the Register contains trials identified from the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, MEDLINE In-Process, MEDLINE Epub Ahead of Print, ClinicalTrials.gov, WHO ICTRP, UK Clinical Research Network Portfolio, and handsearching of journals and conference proceedings. Many of the trials in the Cochrane Incontinence Specialised Register are also contained in CENTRAL.

The date of the last search was 12 February 2018.

The terms used to search the Cochrane Incontinence Specialised Register are given in Appendix 1.

We performed additional searches for the Brief Economic Commentary (BEC) in the following databases:

- MEDLINE on OvidSP (1946 to week 4 July 2018) searched on 2 August 2018;
- Embase on OvidSP (1974 to week 31 2018) searched on 2 August 2018;
- NHS Economic Evaluation Database (NHS EED) on the Centre for Reviews and Dissemination website searched on 3 August 2018.

The search strategies used for the BEC are given in Appendix 2.

### Searching other resources

We cross-referenced relevant conference abstracts identified from the Cochrane Incontinence Specialised Register search to determine if a full-length report had been published. We sought additional trials from the reference lists of included trials.

## Data collection and analysis

### Selection of studies

Two review authors (CD with LPC or JHS) independently screened the list of titles and abstracts generated by our search. We retrieved full-text articles of potentially relevant studies. We also included trials for which only abstracts were available. Two review authors (CD with LPC or JHS) independently assessed the full-text articles or abstracts for eligibility. We contacted study investigators as required. We resolved any differences of opinion by discussion or involvement of a third party. We listed studies formally considered for the review but excluded, with the reasons given for their exclusion. The selection process is documented with a PRISMA flow chart (see Figure 1).

**Figure 1. PRISMA study flow diagram.**

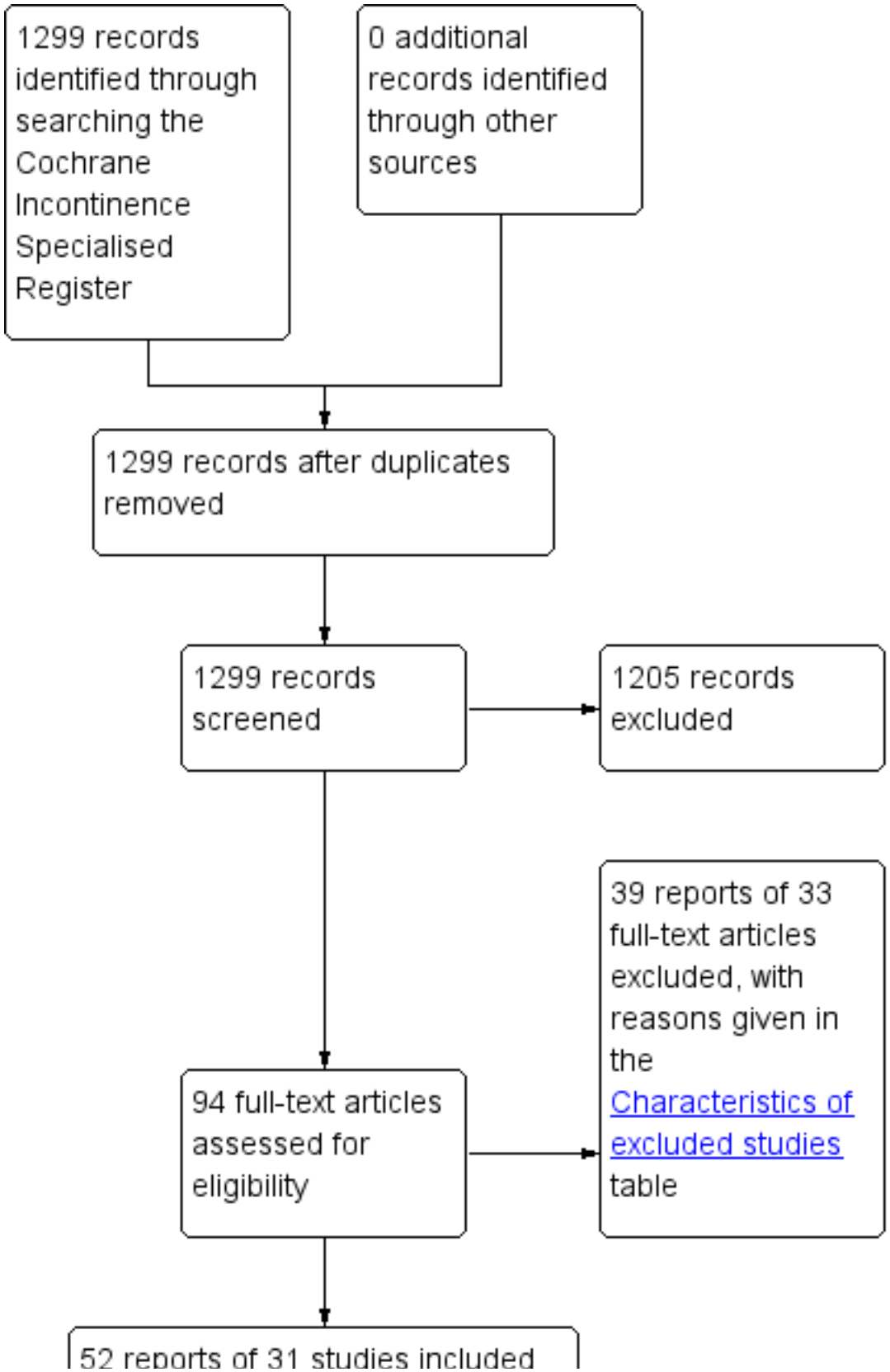
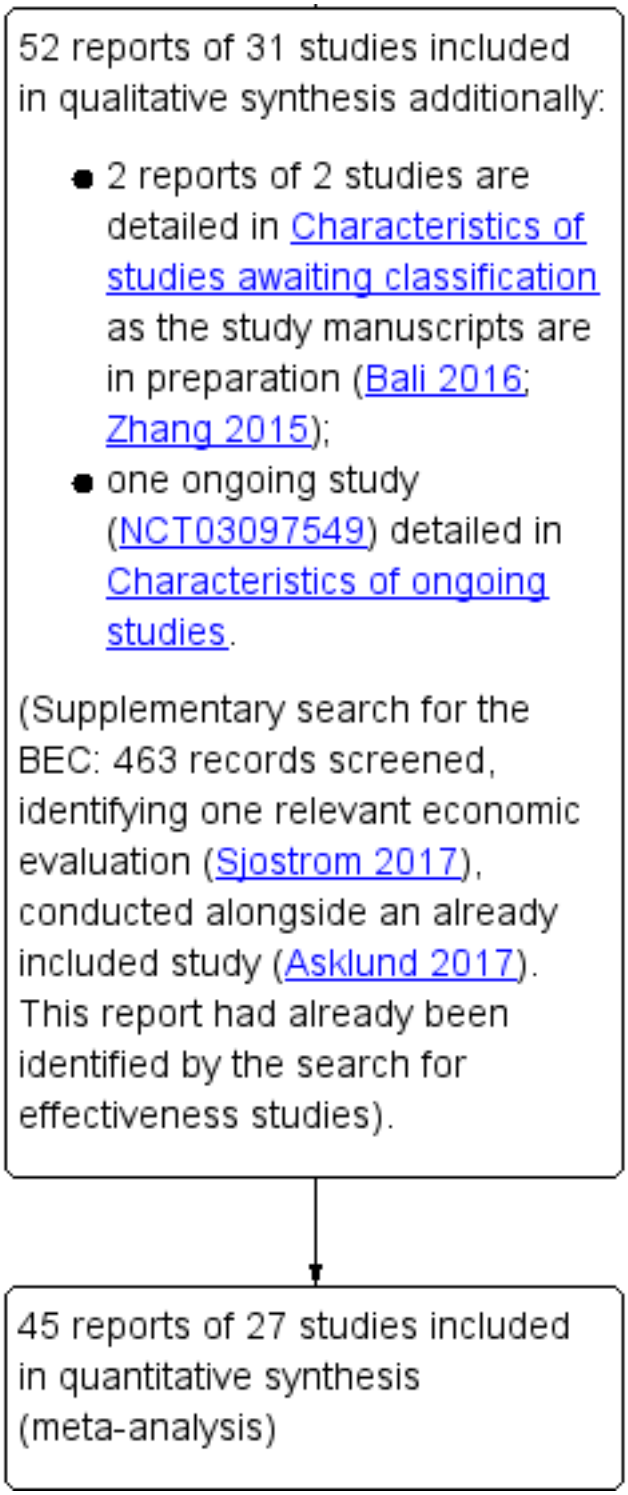


**Figure 1. (Continued)**




**Data extraction and management**

Two review authors (CD and LPC) independently undertook data extraction, which was cross-checked by a third review author (JHS). We resolved any differences of opinion related to the data extraction by discussion. Where study data were possibly collected but not reported, or data were reported in a form that could not be used in the formal comparisons, we sought further clarification from the trialists. In addition, where the reported data were clearly incomplete (i.e. data from abstracts for ongoing trials), we contacted the trialists for data from the completed trial. When found, we added data to the extraction sheet. Two review authors (CD and LPC) performed data entry using Review Manager 5 software (Review Manager 2014). All included trial data were processed as described in the *Cochrane Handbook for Systematic*

*Reviews of Interventions* (Higgins 2011). One review author (JHS) cross-checked data entry. We resolved any differences of opinion related to the data extraction by discussion.

For categorical outcomes, we related the numbers reporting an outcome to the numbers at risk in each group to derive a risk ratio (RR) with 95% confidence intervals (CI). For continuous variables, we used means and standard deviations (SD) to derive mean differences (MD) and 95% CIs. We had planned to undertake formal meta-analysis where appropriate.

### Assessment of risk of bias in included studies

We assessed the risk of bias in the included trials using Cochrane's 'Risk of bias' assessment tool (Higgins 2011). This included the following.

- Random sequence generation (selection bias).
- Allocation concealment (selection bias).
- Blinding of participants and personnel (performance bias).
- Blinding of outcome assessment (detection bias).
- Selective reporting (reporting bias).
- Incomplete outcome data (attrition bias).
- Baseline comparability of the randomised groups.

Baseline comparability was included in the risk of bias assessment because difference in incontinence severity between groups (for example) can influence outcomes. This may be even more important in small trials, where randomisation cannot insure comparability between groups.

Two review authors (CD and LPC) independently assessed these domains, which another review author (JHS) cross-checked. Any differences of opinion were resolved by consensus.

### Measures of treatment effect

Analyses were based on available data from all included trials relevant to the comparisons and outcomes of interest. For trials with multiple publications, we included only the most up-to-date or complete data for each outcome. We undertook meta-analysis where data were available from more than one study assessing the same outcome. We used a fixed-effect model for calculations of pooled estimates and their 95% CIs.

For categorical outcomes, we related the numbers reporting an outcome to the numbers at risk in each group to calculate an RR with 95% CI. For continuous variables, we used means and SDs to calculate an MD with 95% CI. For positive outcomes such as cure, we altered the labelling of the forest plots. If data to calculate RRs or MDs were not given, we utilised the most detailed numerical data available to calculate the actual numbers or means and SDs (e.g. test statistics, P values).

### Unit of analysis issues

The primary analysis was per woman randomised.

### Dealing with missing data

Where possible, we analysed the trial data according to the intention-to-treat principle, that is by the randomised groups, and irrespective of whether women received treatment according to their randomised allocation. We did not impute missing outcome data.

If trials reported sufficient detail to calculate MDs but not enough information to calculate the associated SD, we assumed the outcome to have an SD equal to the highest SD from other trials within the same analysis.

We attempted to obtain missing data from the original trialists.

### Assessment of heterogeneity

We assessed heterogeneity between trials by visual inspection of plots of the data, the $Chi^2$ test for heterogeneity and the $I^2$ statistic (Higgins 2011). We defined the thresholds for interpretation of the $I^2$ statistic according to the *Cochrane Handbook for Systematic Reviews of Interventions*, considering substantial heterogeneity for $I^2$ values between 50% and 90% and considerable heterogeneity for $I^2$ values higher than 75% (Higgins 2011). We sought and discussed possible explanations for heterogeneity.

### Assessment of reporting biases

In view of the difficulty of detecting and correcting for publication bias and other reporting biases, we aimed to minimise their potential impact by ensuring a comprehensive search for eligible studies and by being alert for duplication of data. With respect to reporting bias, and more specifically publication bias, we planned to use the Egger's test where there were more than 10 studies per subgroup analysis.

### Data synthesis

For dichotomous outcomes, we pooled RRs using the Mantel-Haenszel method. For continuous outcomes, we presented MDs using inverse variance. We used a fixed-effect model approach in the meta-analyses in this review except when there was significant heterogeneity ($Chi^2$ test, P less than 0.10), where we used a more conservative random-effects model.

### Subgroup analysis and investigation of heterogeneity

We used subgroup analysis to address the effect of the type of incontinence on outcome. Because the rationale for PFMT is different for the two main types of UI (SUI and UUI), it is plausible to expect a difference in the outcome of PFMT on the basis of the type of incontinence. It is commonly believed that PFMT is most effective for women with SUI and that it may be effective, in combination with behavioural interventions, for women with MUI. In the past, PFMT has rarely been the first-choice treatment for women with UUI alone (Dumoulin 2017).

The four prespecified diagnostic subgroups were trials that recruited women with:

- SUI alone (symptoms, signs, urodynamic stress incontinence);
- UUI alone (symptoms, signs, idiopathic detrusor overactivity incontinence);
- MUI (symptoms or signs of both SUI and UUI, or idiopathic detrusor overactivity incontinence with urodynamic SUI); and
- a range of diagnoses of UI (women could have SUI, UUI or MUI, but data were not reported separately according to these subgroups).

If heterogeneity between trials was sufficiently large, we planned to conduct an investigation to identify its causes. The investigation of heterogeneity addressed the populations and interventions in the individual trials. If heterogeneity remained after appropriate investigation and possible removal of outlying trials, we planned to use a random-effects model in the meta-analysis.

#### Sensitivity analysis

Had sufficient data been available, we planned sensitivity analysis with respect to trial quality, as there is some evidence that the adequacy of randomisation (sequence generation and allocation concealment) may have an impact on the findings of a meta-analysis (Moher 1998).

#### *'Summary of findings' tables and assessment of the quality of the evidence*

We prepared 'Summary of findings' tables for our main comparisons and presented the results for the outcomes prespecified for this purpose in the Types of outcome measures section.

We adopted the GRADE approach to assess the quality of evidence related to these outcomes (Guyatt 2013a; Guyatt 2013b). The four levels of evidence quality were 'high,' 'moderate,' 'low' or 'very low.' The following factors were considered for assessing the quality of evidence: limitations in the study design, inconsistency of results, indirectness of evidence, imprecision and publication bias.

#### *Incorporating economics evidence*

A Brief Economic Commentary was developed to summarise the availability and principal findings of the economic evaluations captured as part of this review. This included evaluations alongside trials and model-based evaluations. This was carried out in accordance with current guidance. This commentary focused on the extent to which principal findings of eligible economic evaluations indicate that an intervention might be judged favourably (or unfavourably) from an economic perspective when implemented in different settings. A supplementary search to identify economic studies was carried out according to the guidelines in Cochrane Economics Methods (Shemilt 2011).

# RESULTS

## Description of studies

### Results of the search

The literature search produced 1299 records for screening, from which we retrieved 94 new potentially relevant full-text articles. There were 52 reports of 31 studies that met the inclusion criteria and we excluded 39 reports of 33 studies, with reasons given in the Characteristics of excluded studies. In addition, two studies were placed in Studies awaiting classification pending further information from the trialists (Bali 2016; Zhang 2015), and one study was placed in Ongoing studies (NCT03097549).

The supplementary search for the BEC produced 463 records which we screened and found one relevant economic evaluation (Sjostrom 2017), conducted alongside an included study (Asklund 2017). This report had already been identified by the search for effectiveness studies.

The flow of literature through the assessment process is shown in the PRISMA flowchart (see Figure 1).

### Included studies

For more details about the trials, see the Characteristics of included studies.

Of the 31 included trials, four contained no data usable in forest plots (Bidmead 2002; Ferreira 2014; Miller 1998; Wells 1999), while 27 contributed to forest plots. Twenty-one trials contributed to the analysis of primary outcomes:

- cure (Bø 1999; Burgio 1998; Hofbauer 1990; Kim 2007; Kim 2011a; Kim 2011b);
- cure or improvement (Asklund 2017; Bø 1999; Burgio 1998; Diokno 2010; Lagro-Janssen 1991a; Leong 2015);
- symptom- or condition-specific health measures (Asklund 2017; Bertotto 2017; Beuttenmuller 2010; Bø 1999; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Firra 2013; Kargar Jahromi 2013; Leong 2015; McLean 2013; Pereira 2011; Solberg 2016; Sran 2016).

Seventeen trials had more than two treatment arms (Aksac 2003; Bertotto 2017; Beuttenmuller 2010; Bidmead 2002; Bø 1999; Burgio 1998; Burns 1993; Castro 2008; Firra 2013; Henalla 1989; Henalla 1990; Hofbauer 1990; Kim 2011a; Pereira 2011; Solberg 2016; Wells 1999; Yoon 1999). Although all comparison groups are listed (see Characteristics of included studies), only descriptions and data relating to the PFMT and control arms were given in this review. Of the 31 included trials, 21 were included in the previous version of the review (Dumoulin 2014).

#### *Design*

All included trials were RCTs except one, which was considered to be quasi-randomised (Lagro-Janssen 1991a).

#### *Sample sizes*

Sample size ranged from 15 to 143 participants per study.

#### *Setting*

The settings were single centre in 22 trials (Aksac 2003; Bertotto 2017; Beuttenmuller 2010; Bidmead 2002; Burgio 1998; Burns 1993; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Ferreira 2014; Firra 2013; Henalla 1989; Henalla 1990; Hofbauer 1990; Kargar Jahromi 2013; McLean 2013; Miller 1998; Pereira 2011; Sar 2009; Solberg 2016; Sran 2016; Yoon 2003), or multi-centre in four trials (Bø 1999; Lagro-Janssen 1991a; Leong 2015; Wells 1999). In three other trials, participants came from a multiples resident register in the USA (Diokno 2010); a single resident register in Tokyo, Japan (Kim 2007; Kim 2011a; Kim 2011b); or through a website (Asklund 2017). Fourteen countries contributed studies to this review: USA (six); Brazil (five); UK, Japan and Turkey (three each); Canada and Norway (two each); and Austria, China, Iran, Korea, Portugal, The Netherlands and Sweden (one each).

#### *Participants*

All the women in the included studies had UI. Fifteen trials diagnosed the type of UI based on symptoms or signs, or both. Of these:

- seven included only women with SUI (Asklund 2017; Bertotto 2017; Ferreira 2014; Kargar Jahromi 2013; Kim 2007; Miller 1998; Pereira 2011);
- seven included women with all types of UI (Diokno 2010; Kim 2011a; Kim 2011b; Leong 2015; Sar 2009; Sran 2016; Yoon 2003); and
- one included women with SUI or MUI (Wells 1999).

Thirteen trials reported urodynamic diagnoses:

- eight included women with urodynamic SUI only (Aksac 2003; Bidmead 2002; Bø 1999; Carneiro 2010; Castro 2008; Henalla 1989; Henalla 1990; Hofbauer 1990);
- one included women with detrusor overactivity incontinence with or without urodynamic SUI and 51% had MUI (Burgio 1998);
- one included women with urodynamic SUI with or without detrusor overactivity incontinence, but the proportion with mixed symptoms was small (9%) (Burns 1993);
- one included women with urodynamic SUI and MUI, with 54% of the participants having mixed symptoms (Celiker Tosun 2015);
- one included women with SUI, UUI or MUI, although a subset of data was available for women with urodynamic SUI only (Lagro-Janssen 1991a); and
- one included women with symptoms of SUI or MUI and excluded those with detrusor instability (McLean 2013).

For the other three trials, one included women with all types of UI diagnosed by symptoms or urodynamics (Firra 2013), and two were unclear about the diagnosis method for MUI (Beuttenmuller 2010), and SUI (Solberg 2016).

Based on diagnosis, the incontinence subgroups used in the analysis were:

- SUI: 18 trials (Aksac 2003; Asklund 2017; Bertotto 2017; Beuttenmuller 2010; Bø 1999; Burns 1993; Carneiro 2010; Castro 2008; Firra 2013; Henalla 1989; Henalla 1990; Hofbauer 1990; Kargar Jahromi 2013; Kim 2007; Kim 2011a; Lagro-Janssen 1991a; McLean 2013; Pereira 2011);
- MUI: one trial (Solberg 2016);
- UUI: one trial (Firra 2013); and
- UI, range of diagnoses: nine trials (Burgio 1998 (UUI/MUI); Celiker Tosun 2015 (SUI/MUI); Diokno 2010 (all types of UI); Kim 2011a (all types of UI); Kim 2011b (all types of UI); Leong 2015 (all types of UI); Sar 2009 (all types of UI); Sran 2016 (all types of UI); Yoon 2003 (all types of UI)).

Lagro-Janssen 1991a recruited women with SUI, UUI or MUI, and those with UUI or MUI were offered bladder training. As data from women with SUI were reported separately, this trial was eligible for inclusion.

For two trials, women with a range of diagnoses were recruited but data was reported by subgroups of symptoms. In one case, there were data available separately for women with predominant SUI or UUI symptoms (Firra 2013). In the other trial, there were data reported for SUI and UUI or MUI together (Kim 2011a).

### *Other characteristics*

In 12 trials, leakage frequency was one of the inclusion criteria:

- more than once a month (Kim 2007; Kim 2011a; Pereira 2011);
- twice or more per month (Lagro-Janssen 1991a);
- once or more per week (Asklund 2017; Kim 2011b);
- twice or more per week (Burgio 1998);
- three times or more per week (Burns 1993; Castro 2008);
- twice or more per three days (Sran 2016);
- one to five leakage episodes per day (Miller 1998); and
- three incontinence episodes or more in three days (Firra 2013).

Four trials used amount of leakage from a pad test:

- more than 1 g during 15 minutes of volleyball practice (Ferreira 2014);
- more than 1 g during a 30-minute test (Yoon 2003);
- more than 2 g during a 60-minute pad test (Sar 2009); and
- more than 4 g on a short clinic-based pad test, with standardised bladder volume (Bø 1999).

Aside from diagnosis and some measure of leakage severity, there were no other consistently reported inclusion criteria, although 18 trials restricted participation based on age. These trials recruited women aged:

- 13 to 30 years (Ferreira 2014);
- 18 years and older (Asklund 2017; McLean 2013; Solberg 2016);
- 20 to 65 years (Lagro-Janssen 1991a);
- 21 years and older (Firra 2013);
- 35 to 50 years (Carneiro 2010);
- 35 to 55 years (Yoon 2003);
- 50 to 65 years (Bertotto 2017);
- 55 years and older (Burgio 1998; Burns 1993; Sran 2016);
- 60 years and older (Miller 1998);
- 60 to 74 years (Kargar Jahromi 2013);
- 65 years and older (Leong 2015); and
- 70 years and older (Kim 2007; Kim 2011a; Kim 2011b).

Common exclusion criteria were untreated urinary tract infection, postmicturition residual volume greater than a specified amount, neurological disorders and cognitive impairment. Three trials also considered body mass index (BMI) as a criteria for exclusion when it was greater than 30 kg/m$^2$ (Leong 2015), greater than 50 kg/m$^2$ (Firra 2013), or either lower than 18 kg/m$^2$ or greater than 25 kg/m$^2$ (Ferreira 2014).

### *Interventions*

The individual characteristics of the active and control interventions are summarised in the Characteristics of included studies and detailed in Appendix 3.

#### Active intervention: pelvic floor muscle training

Three trials gave no details of the PFMT programme used (Bidmead 2002; Henalla 1990; Hofbauer 1990). Of the 28 remaining trials, 21 stated that a correct VPFMC was confirmed prior to training using vaginal, rectal or physical examination (Aksac 2003; Bertotto 2017; Bø 1999; Burgio 1998; Burns 1993; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Diokno 2010; Firra 2013; Henalla 1989; Lagro-Janssen 1991a; Leong 2015; McLean 2013; Miller 1998; Pereira 2011; Sar 2009; Solberg 2016; Sran 2016; Wells 1999; Yoon 2003). Five

trials reported that participants were taught to do a VPFMC but did not say how they were taught (Ferreira 2014; Kargar Jahromi 2013; Kim 2007; Kim 2011a; Kim 2011b). One trial did not report correct VPFMC confirmation, but stated that the pelvic floor muscles were assessed by an evaluator prior to treatment (Beuttenmuller 2010). Another trial reported that participants were instructed by a smartphone app to identify the correct VPFMC, but without face-to-face interaction with health professionals (Asklund 2017).

Specialist nurses taught PFMT in 10 trials (Burgio 1998; Burns 1993; Diokno 2010; Kim 2007; Kim 2011a; Kim 2011b; Miller 1998; Sar 2009; Wells 1999; Yoon 2003), physiotherapists in 16 trials (Bertotto 2017; Beuttenmuller 2010; Bidmead 2002; Bø 1999; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Firra 2013; Henalla 1989; Henalla 1990; Hofbauer 1990; Leong 2015; McLean 2013; Pereira 2011; Solberg 2016; Sran 2016), family doctor in one trial (Lagro-Janssen 1991a), smartphone app in one trial (Asklund 2017), and was not specified in three trials (Aksac 2003; Ferreira 2014; Kargar Jahromi 2013).

Based on the descriptions of training, two trials had PFMT programmes that clearly or predominantly targeted co-ordination or strength training (Bø 1999; Miller 1998). Miller 1998 described a short (one-week) programme to improve co-ordination between a VPFMC and a rise in intra-abdominal pressure. Bø 1999 recommended a programme that comprised of 8 to 12 high-intensity (close to maximal) VPFMC, with six to eight second hold and three to four fast contractions added at the end of each hold, and six-second rests between contractions, three times per day. Exercises were done in different body positions, including lying, kneeling, sitting and standing, all with legs apart.

It was more difficult to characterise or categorise the other PFMT programmes because they were either a mixed programme (e.g. strength and endurance) or had not described a key training parameter (e.g. amount of voluntary effort per contraction). The individual characteristics of each exercise programme (i.e. the number of VPFMC, duration of holding time, duration of rest time, number of sets per day, types of contraction strength, endurance, co-ordination, body position and adherence strategies) are detailed in Appendix 3.

Many of the recent trials described a mixed programme of short or short and rapid contractions of one to three seconds and long sustained contractions of 6 to 59 seconds (Asklund 2017; Celiker Tosun 2015; Diokno 2010; Ferreira 2014; Firra 2013; Kargar Jahromi 2013; Kim 2011a; Kim 2011b; Leong 2015; Sar 2009; Sran 2016), in addition to contraction prior to and during a cough (Asklund 2017; Castro 2008; Celiker Tosun 2015; Diokno 2010; Leong 2015; McLean 2013; Sar 2009; Sran 2016), or prior to an abdominal strain (Bertotto 2017), and in different body positions from lying to standing (Asklund 2017; Bertotto 2017; Beuttenmuller 2010; Carneiro 2010; Firra 2013; Kargar Jahromi 2013; Kim 2007; Kim 2011a; Kim 2011b; Leong 2015; Pereira 2011; Sar 2009; Sran 2016).

The training programme was progressive in 14 trials, increasing the difficulty of the exercise week by week, including body position or number of repetitions (Asklund 2017; Bertotto 2017; Burns 1993; Carneiro 2010; Celiker Tosun 2015; Firra 2013; Leong 2015; Pereira 2011; Sran 2016), or the contraction holding time (Aksac 2003; Bø 1999; Burgio 1998; Sar 2009; Yoon 2003).

#### Control interventions

Control interventions included:

- no treatment (Aksac 2003; Asklund 2017; Bertotto 2017; Beuttenmuller 2010; Bidmead 2002; Burns 1993; Carneiro 2010; Celiker Tosun 2015; Diokno 2010; Firra 2013; Henalla 1989; Henalla 1990; Kargar Jahromi 2013; McLean 2013; Miller 1998; Pereira 2011; Sar 2009; Solberg 2016; Yoon 2003);
- placebo drug (Burgio 1998);
- sham electrical stimulation (Hofbauer 1990); and
- other inactive control treatments that comprised:
  - use of an anti-incontinence device (Bø 1999);
  - advice on incontinence pads (Lagro-Janssen 1991a);
  - motivational telephone calls once per month (Castro 2008);
  - advice on simple lifestyle alterations (Kim 2011b; Wells 1999);
  - general education class (cognitive function, osteoporosis and oral hygiene) (Kim 2011a; Sran 2016);
  - refraining from special exercises aiming to increase muscle strength, walking speed, to reduce BMI or to improve dietary habits (Kim 2007); and
  - access to an educational pamphlet with or without advice on UI (Ferreira 2014; Leong 2015).

### *Outcomes*

Overall, there was no consistency in the choice of outcome measures by trialists. This limited the possibilities for considering the results from individual trials together. It was disappointing that four eligible trials did not contribute any data to the main analyses because they did not report any prespecified outcomes of interest, or they did not report their outcome data in a usable way (e.g. mean without a measure of dispersion, P values without raw data, or only postintervention minus pre-intervention data or differential data available) (Bidmead 2002; Ferreira 2014; Miller 1998; Wells 1999). We attempted communication with the authors but received no responses.

As the length of intervention and timing of postintervention assessment varied, there was no attempt to report outcomes at a particular time point. Postintervention outcomes were used as it was assumed that the trialists would choose to complete treatment and measure outcomes when maximum benefit was likely to have been gained. Data from after treatment stopped or any longer-term follow-up were reported as secondary outcomes.

For one trial, the length of the programme depended on the strength of the participants' pelvic floor muscles (Celiker Tosun 2015). Those who did not reach the goal of grade 5 (according to the Oxford grading system) at the end of the 12-week PFMT programme were invited for additional training until the goal was achieved. However, only data from the primary endpoint (after the 12-week training programme) was eligible for this review since there was no additional follow-up for the control group.

### *Primary outcomes – participant-reported measures*

#### Symptomatic cure or symptomatic cure or improvement of urinary incontinence at the end of treatment

The studies used many different scales to measure a participant's response to treatment, including Likert scales, visual analogue scales and per cent reduction in symptoms. Whatever the scale,

data were included in the formal comparisons when the trialists stated the number of women who perceived they were cured or improved, as defined by the trialists, after treatment. Where there was more than one level of improvement reported (e.g. much better and somewhat better), we entered data for the greater degree of improvement in the comparison. It was thought that this was more likely to capture participants who had clinically important improvement. As some trial reports did not differentiate cure from improvement, we used two measures to avoid losing important data ('cure only' or 'cure or improvement').

The studies used the following definitions.

- Participant perceived cure defined as no urine loss or 'dry' (Burgio 1998; Kim 2011a).
- Participant perceived cure as 'incontinence is now unproblematic' (Bø 1999).
- Cure was also reported by women as no leakage in a urinary diary (Hofbauer 1990; Kim 2007; Kim 2011b).
- Participant perceived cure and improvement defined as much better and somewhat better (Asklund 2017; Diokno 2010).
- Participant perceived cure and improvement defined as '75% or more perceived improvement' (Burgio 1998).
- Participant perceived cure and improvement defined as 'dry' or 'improved' (Lagro-Janssen 1991a).
- Participant perceived cure and improvement defined as 'continent' or 'almost continent' (Bø 1999).

#### Patient Global Impression of Improvement

Asklund 2017 used the Patient Global Impression of Improvement (PGI-I) questionnaire, which is highly recommended to measure symptom bother related to
UI in women with SUI (Grade A, Kelleher 2013). It is a validated questionnaire that asks about the change experienced after treatment, with seven response options ranging from "very much better" to "very much worse" (Yalcin 2003).

Leong 2015 used a visual analogue scale (0 suggesting "no improvement" and 10 "complete relief") to measure participant perception of improvement (Analysis 1.18).

#### Symptom- and condition-specific quality of life measures

Thirteen trials used psychometrically robust questionnaires (Grade A questionnaire according to International Consultation on Incontinence book) (Kelleher 2013) for assessment of incontinence symptoms, or the impact of these symptoms on QoL, or both. They are presented below and can be found in the forest plots.

#### King's Health questionnaire

Three trials used the King's Heath questionnaire (Beuttenmuller 2010; Carneiro 2010; Pereira 2011), which has established validity, reliability and responsiveness to change or evaluation of UI symptoms in women (Kelleher 1997; Margolis 2011; Grade A+, Kelleher 2013). Higher scores represent higher QoL. Four domains of this questionnaire (severity measure, incontinence impact, physical limitation and general health scores) are presented in the forest plots (Analysis 1.3; Analysis 1.4; Analysis 1.5; Analysis 1.6).

#### International Consultation on Incontinence Modular Questionnaire – Short Form

Four trials used the International Consultation on Incontinence Modular Questionnaire – Short Form (ICIQ-UI Short Form) (Asklund 2017; Bertotto 2017; Kargar Jahromi 2013; Solberg 2016), which is highly recommended for the assessment of the impact of UI on QoL (Grade A, Kelleher 2013). It includes four items: frequency of UI, amount of leakage, overall impact of UI and self-diagnosis. The score ranges from 0 to 21, with greater values indicating increased severity (Avery 2004).

Asklund 2017 used the lower urinary tract symptoms module of the ICIQ questionnaire (ICIQ-LUTSqol), which is also highly recommended to assess the impact of lower urinary tract symptoms on QoL with particular reference to social effects (Grade A+, Kelleher 2013). It is a 19-item questionnaire, with scores ranging from 19 to 76 and higher values indicating lower QoL (Brookes 2004).

#### Incontinence Impact Questionnaire

Three trials used the short form version of the Incontinence Impact Questionnaire (IIQ) (Celiker Tosun 2015; Leong 2015; McLean 2013), and one used the long form version (Sran 2016). Both versions are highly recommended for the assessment of QoL impact of UI (Grade A, Kelleher 2013). The score ranges from 0 to 400 for the long form (Wyman 1987), and from 0 to 100 for the short form, with higher scores indicating worse symptoms (Uebersax 1995).

#### Urogenital Distress Inventory

Three trials used the short form of the Urogenital Distress Inventory (UDI) questionnaire (UDI-6) (Celiker Tosun 2015; McLean 2013; Sran 2016), which assesses the degree to which symptoms associated to incontinence are troubling. The long and short versions are highly recommended for the assessment of symptoms of UI in women (Grade A, Kelleher 2013). The short form includes six questions with total scores ranging from 0 to 100, with greater values indicating worse symptoms (Uebersax 1995).

#### Incontinence-specific quality of life

Castro 2008 and Sar 2009 used the UI-specific QoL instrument (I-QOL), which has established validity, reliability and responsiveness to change or evaluation of incontinence symptoms in women (Bushnell 2005; Wagner 1996; Grade A+, Kelleher 2013). The I-QOL contains 22 items, each with a 5-point Likert-type response scale yielding a total score of 0 to 100, with the higher scores representing greater QoL. Castro 2008 reported the total score after treatment, while Sar 2009 only reported change from baseline (thus a positive value for change from pre- to post-treatment represents a deterioration in QoL).

Lower-grade symptoms and condition-specific QoL measures are listed below and are presented in Appendix 4 and Appendix 5.

- B-FLUTS: used by Bø 1999.
- The Social Activity Index: Bø 1999 reported a symptom score that addressed activity limitation (difficulty with certain activities and functions) in nine social situations.
- Sandvik's Incontinence Severity Index (ISI) for female UI: reported by Diokno 2010.
- York Incontinence Perceptions Scale (YIPS): used by Firra 2013.
- Leakage Index: reported by Bø 1999.

- Urine leakage score: Kim 2011b reported a urine leakage score based on a self-reported one week urinary diary. There was no information on the psychometric properties of this instrument.
- UI score: Yoon 2003 reported on a UI score calculated from a 5-point Likert-type scale regarding severity of leakage with 18 prespecified activities associated with urine loss. There was no information on the psychometric properties of this instrument.

### *Secondary outcomes – participant-reported measures*

#### Longer-term symptomatic cure and improvement (six months to one year after end of treatment; more than one year after end of treatment)

Most of the trials evaluated cure or cure and improvement immediately after the treatment period. Only two trials evaluated cure in the medium-term, which was nine months (Henalla 1989) and seven months (Kim 2011b) after treatment.

No trials evaluated cure or improvement one year or more after the end of treatment.

One trial evaluated symptoms and UI-specific QoL using UDI-6 and IIQ questionnaires nine months after the end of treatment (Sran 2016).

#### Satisfaction and need for further treatment

Three trials reported on participant-perceived satisfaction following the intervention (Bø 1999; Burgio 1998; Castro 2008), while two reported on the number of women needing further treatment (Bø 1999; Burgio 1998).

#### Self-efficacy

One trial reported UI-specific self-efficacy using the Geriatric Self-Efficacy Index for Urinary Incontinence (GSE-UI), which assesses self-efficacy to prevent unwanted urine loss among postmenopausal women (Sran 2016). It contains 12 questions proposed for geriatric women with UI, with scores ranging from 0 to 120 and higher scores indicating higher self-efficacy. It has established validity, reliability and responsiveness to change in elderly women with UI (Tannenbaum 2009; Appendix 6). One trial reported on the need for the use of aids (Asklund 2017; Appendix 7).

### *Participant-reported quantification of symptoms*

#### Number of leakage episodes (in 24 hours)

Fourteen of the trials used diaries to collect data on leakage episodes for:

- two days (Yoon 2003);
- three days (Bø 1999; Celiker Tosun 2015; Firra 2013; McLean 2013; Sar 2009);
- four days (Wells 1999);
- seven days (Asklund 2017; Castro 2008; Lagro-Janssen 1991a; Leong 2015; Sran 2016); and
- 14 days (Burgio 1998; Burns 1993).

Yoon 2003 collected these data but did not report them directly. Rather, they reported leakage per 48 hours as an incontinence score. Sar 2009 reported mean change from baseline, while Wells 1999 reported means without a measure of dispersion. To enable comparison between trials, the data were presented as number of leakage episodes in 24 hours.

#### Number of micturitions during the day (frequency) or during the night (nocturia)

Four trials reported on number of micturition during the day (Celiker Tosun 2015; Diokno 2010; Firra 2013; Yoon 2003), while three trials reported on frequency per night (Celiker Tosun 2015; Diokno 2010; Yoon 2003).

### *Clinicians' measures*

#### Pad and paper towel testing in a short test (up to one hour) or long test (24 hours) (grams of urine lost) and number cured or improved based on pad weights in short clinic-based pad test

Thirteen trials reported data on pad and paper towel tests.

- Ten trials used clinic-based short pad tests (Aksac 2003; Bidmead 2002; Bø 1999; Castro 2008; Celiker Tosun 2015; Henalla 1989; Henalla 1990; McLean 2013; Pereira 2011; Yoon 2003).
- In addition to the short pad test, one trial used a 24-hour home-based pad test (Bø 1999).
- One trial used a paper towel test (Miller 1998).
- Two further trials reported only a 24-hour pad test (Diokno 2010; Sran 2016).

Aside from differences in the type of test, trialists also presented their data differently. Data were usually categorised (such as cured, improved, not improved) or reported as a mean with SD. The former data were used to report the number of women with objective cure or improvement of incontinence, while the latter were reported as grams of urine lost.

### *Quality of life (not condition-specific)*

#### General health status measures

Two trials reported non-condition-specific QoL data (Bø 1999; Burgio 1998). Burgio 1998 used the Hopkins Symptom Checklist for psychological distress with 90 items and a total score (Global Severity Index) (Derogatis 1983). Bø 1999 used the Norwegian version of the Quality of Life Scale (QoLS-N) to assess general health and QoL prior to and after the intervention (Wahl 1998).

One trial used the Rosenberg self-esteem questionnaire (Kargar Jahromi 2013).

For further details, see Appendix 6.

#### Sexual function or problems

One trial reported the effect of PFMT on UI during intercourse and in terms of interference with sexual satisfaction (Bø 1999).

### *Adverse effects*

Seven trials reported on adverse effects (Bø 1999; Burgio 1998; Castro 2008; Kargar Jahromi 2013; Lagro-Janssen 1991a; Solberg 2016; Sran 2016).

### *Measure of likely moderator variables*

#### Measures of pelvic floor muscle function

- Three trials used ultrasound (Carneiro 2010; Celiker Tosun 2015; McLean 2013).

- Seven trials used perineometry to measure vaginal squeeze pressure (Aksac 2003; Beuttenmuller 2010; Bø 1999; Celiker Tosun 2015; Firra 2013; Pereira 2011; Yoon 2003).
- Nine trials used digital palpation (Aksac 2003; Beuttenmuller 2010; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Diokno 2010; Miller 1998; Pereira 2011; Wells 1999).
- Four trials used vaginal electromyography (Bertotto 2017; Burns 1993; Carneiro 2010; Wells 1999).

#### Measures of adherence

Twelve trials attempted to measure adherence to PFMT at home using either exercise or training diaries (Asklund 2017; Bidmead 2002; Bø 1999; Burns 1993; Castro 2008; Kim 2007; Kim 2011b; Leong 2015; Solberg 2016; Sran 2016; Wells 1999), or self-reported adherence (Lagro-Janssen 1991a). Five trials attempted to measure attendance at exercise sessions (Kim 2007; Kim 2011b; Leong 2015; Solberg 2016; Sran 2016), with only one study reporting attendance on the control group (for education sessions) (Sran 2016). Results are detailed in Appendix 3.

#### Excluded studies

Full details of the 33 excluded studies are given in the Characteristics of excluded studies.

### Risk of bias in included studies

Figure 2 and Figure 3 summarise the results of the risk of bias analysis.

**Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.**

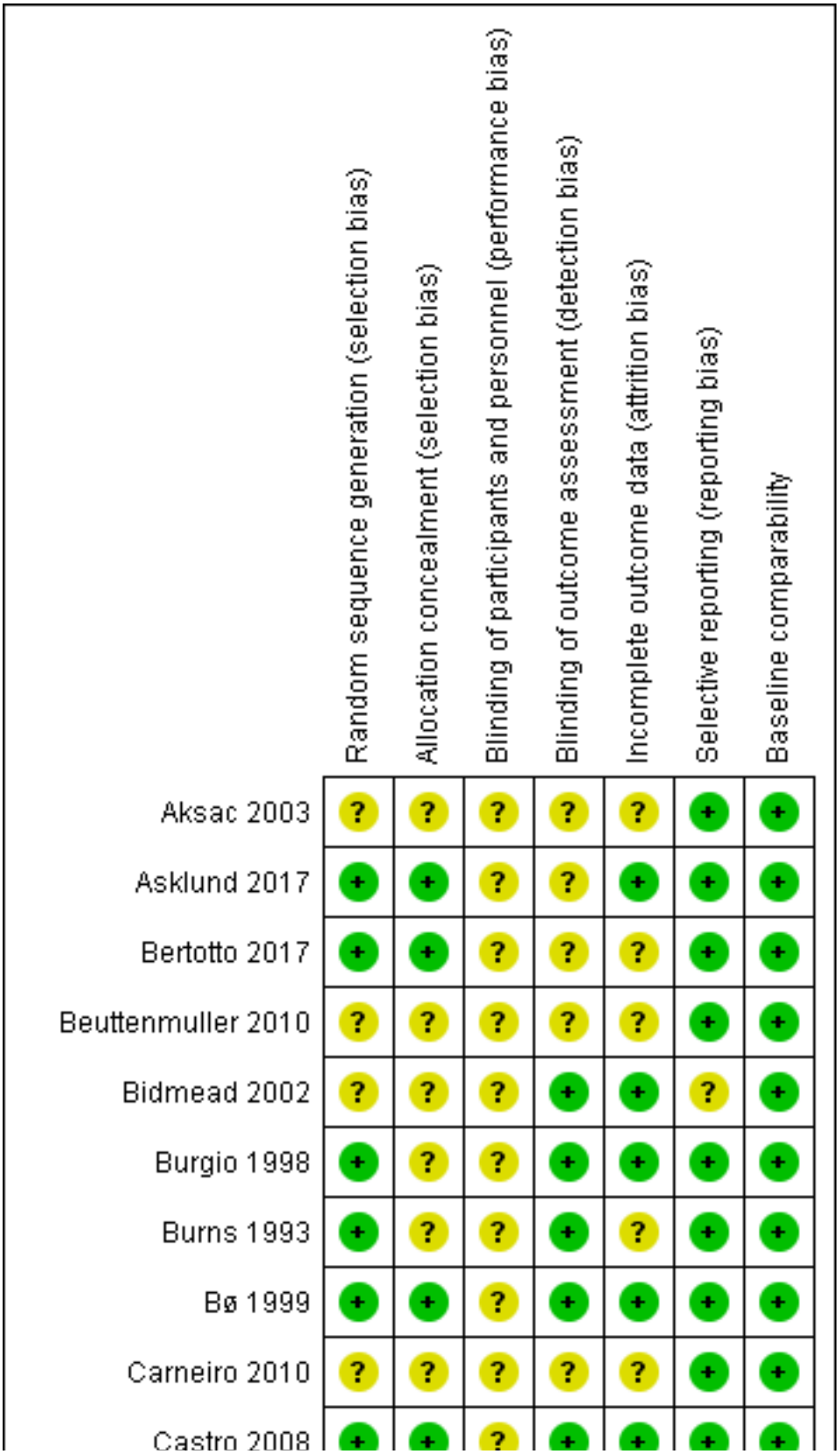
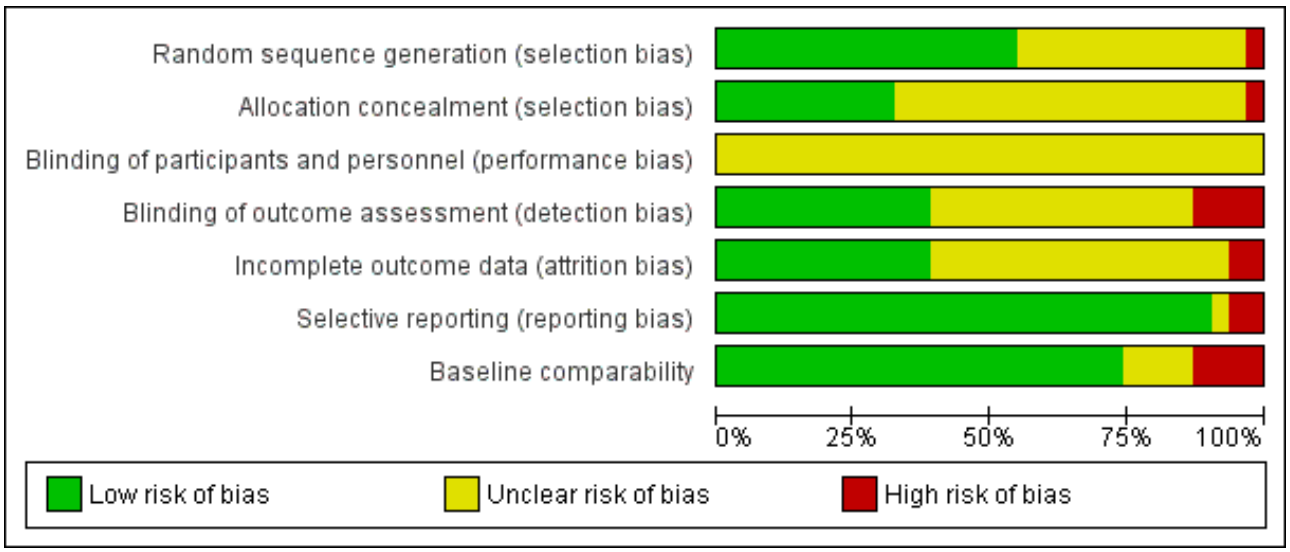


**Figure 2. (Continued)**

**Figure 2. (Continued)**


**Figure 3. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.**




Due to the brevity of reporting, it was difficult to assess the two trials that were published as conference abstracts (Bidmead 2002; Henalla 1990). Fifteen of the trials were small, with fewer than 25 women per comparison group (Aksac 2003; Bertotto 2017; Diokno 2010; Ferreira 2014; Firra 2013; Henalla 1990; Hofbauer 1990; Kargar Jahromi 2013; McLean 2013; Miller 1998; Pereira 2011; Sar 2009; Solberg 2016; Sran 2016; Yoon 2003). Ten were of moderate size, with around 25 to 50 per group (Beuttenmuller 2010; Bø 1999; Burns 1993; Carneiro 2010; Castro 2008; Henalla 1989; Kim 2007; Kim 2011a; Lagro-Janssen 1991a; Leong 2015). The other five allocated more than 50 women per group (Asklund 2017; Burgio 1998; Celiker Tosun 2015; Kim 2011b; Wells 1999). One trial randomised participants in a 2:1 ratio, with 40 in the PFMT group and 20 in the control group (Bidmead 2002). There were no large or very large trials. Eleven trials reported on a priori power calculation (Asklund 2017; Bø 1999; Castro 2008; Celiker Tosun 2015; Firra 2013; Kim 2007; Kim 2011b; Leong 2015; McLean 2013; Sar 2009; Sran 2016).

### Allocation

#### *Random sequence generation*

Seventeen trials generated a genuine random sequence (e.g. computer generation of random numbers, block size) (Asklund 2017; Bertotto 2017; Bø 1999; Burgio 1998; Burns 1993; Castro 2008; Celiker Tosun 2015; Diokno 2010; Ferreira 2014; Kim 2007; Kim 2011a; Kim 2011b; Leong 2015; McLean 2013; Miller 1998; Solberg 2016; Sran 2016). Thirteen trials stated only that women were allocated at random, with no further description (Aksac 2003; Beuttenmuller 2010; Bidmead 2002; Carneiro 2010; Firra 2013; Henalla 1989; Henalla 1990; Hofbauer 1990; Kargar Jahromi 2013; Pereira 2011; Sar 2009; Wells 1999; Yoon 2003).

The abstract of one study stated that women were randomly allocated to comparison groups, but the methods section of the same paper reported that women were "consecutively assigned" (Lagro-Janssen 1991a). Therefore, it appears this was a quasi-randomised trial rather than a randomised trial and was at high risk of bias for this domain.

#### *Allocation concealment*

Ten trials reported allocation concealment adequately (Asklund 2017; Bertotto 2017; Bø 1999; Castro 2008; Celiker Tosun 2015; Kim 2011a; Kim 2011b; Leong 2015; Solberg 2016; Sran 2016). For 20 trials, insufficient information meant it was not clear if allocation was adequately concealed (Aksac 2003; Beuttenmuller 2010; Bidmead 2002; Burgio 1998; Burns 1993; Carneiro 2010; Diokno 2010; Ferreira 2014; Firra 2013; Henalla 1989; Henalla 1990; Hofbauer 1990; Kargar Jahromi 2013; Kim 2007; McLean 2013; Miller 1998; Pereira 2011; Sar 2009; Wells 1999; Yoon 2003). One trial had inadequate allocation concealment (alternate allocation), which was considered to be quasi-randomised and at high risk of allocation concealment (Lagro-Janssen 1991a).

### Blinding

#### *Blinding of participants and personnel (performance bias)*

Given the nature of PFMT it is difficult, and often impossible, to blind the treatment provider and participants during treatment. We rated all trials as unclear for this domain as it is not feasible to blind the participants or care providers. We specified in the 'Risk of bias' tables any in inactive interventions given to the control group in an effort to reduce performance bias.

#### *Blinding of outcome assessment (detection bias)*

Twelve trials reported using blinded outcome assessors (Bidmead 2002; Bø 1999; Burgio 1998; Burns 1993; Castro 2008; Celiker Tosun 2015; Diokno 2010; Kim 2011b; Lagro-Janssen 1991a; Miller 1998; Sran 2016; Yoon 2003).

In 15 trials, the authors did not report sufficient information to conclude that the outcome assessment was blinded (Aksac 2003; Asklund 2017; Bertotto 2017; Beuttenmuller 2010; Carneiro 2010; Ferreira 2014; Henalla 1989; Henalla 1990; Hofbauer 1990; Kargar Jahromi 2013; Kim 2007; Kim 2011a; McLean 2013; Solberg 2016; Wells 1999).

The last four trials reported that the outcome assessors were not blinded to treatment assignment (Firra 2013; Leong 2015; Pereira 2011; Sar 2009).

### Incomplete outcome data

There were no dropouts or loss to follow-up in three trials (Ferreira 2014; Leong 2015; Miller 1998). In six trials it appeared there were no dropouts, but this was not clearly stated in the trial reports (Aksac 2003; Beuttenmuller 2010; Carneiro 2010; Henalla 1989; Henalla 1990; Hofbauer 1990). Twenty-two trials reported attrition, dropouts or losses to follow-up. In these trials the proportion was:

- less than 10% in five (Asklund 2017; Burns 1993; Celiker Tosun 2015; Kargar Jahromi 2013; Kim 2007; Kim 2011a; Kim 2011b; Lagro-Janssen 1991a; Sran 2016);
- 11% to 15% in six (Bertotto 2017; Bø 1999; Burgio 1998; Castro 2008; Diokno 2010; Firra 2013; McLean 2013; Pereira 2011; Yoon 2003);
- more than 20% in two (Bidmead 2002; Sar 2009); and
- nearly 50% in two others (Solberg 2016; Wells 1999).

The proportion of withdrawals or loss to follow-up was higher in the control group in five trials (Bertotto 2017; Burgio 1998; Firra 2013; Sar 2009; Solberg 2016), and in the intervention group in one trial (Celiker Tosun 2015). There were no clear differences in the other trials. In most trials the cause of the differential dropout was not thought to be significantly related to the intervention, but for one case there was differential dropout from the groups (Sar 2009). Five of 22 women were excluded from the control group analysis as they received other treatment for their incontinence and this was not reflected in the analysis of the remaining 17 women.

### Selective reporting

It was unclear if there was selective reporting of the outcomes in most trials because the protocols were not available for most of the older studies. Therefore, we considered risk of bias to be low when it was clear that the published report included all expected outcomes. Two trials did not report all expected outcomes in the results section (Ferreira 2014; Henalla 1990). One trial reported all expected outcomes narratively in the results section but without any supporting data (Bidmead 2002).

### Other potential sources of bias

#### *Baseline comparability*

Twenty-three trials were comparable at baseline for all important outcomes and demographic characteristics that might predict outcomes such as symptom severity or duration (Aksac 2003; Asklund 2017; Bertotto 2017; Beuttenmuller 2010; Bidmead 2002; Bø 1999; Burgio 1998; Burns 1993; Carneiro 2010; Castro 2008; Ferreira 2014; Henalla 1989; Kargar Jahromi 2013; Kim 2007; Kim 2011a; Kim 2011b; Lagro-Janssen 1991a; McLean 2013; Miller 1998; Pereira 2011; Sar 2009; Sran 2016; Yoon 2003). Three trials did not give enough information to assess baseline comparability between groups (Henalla 1990; Hofbauer 1990; Wells 1999). Finally, five trials reported a statistically significant difference between the PFMT and control groups (Celiker Tosun 2015; Diokno 2010; Firra 2013; Leong 2015; Solberg 2016). Differences were reported for age, with the PFMT group being older than the control group in one trial (Diokno 2010), and marginally younger in another (Leong 2015, P = 0.06); number of micturitions during the day (Leong 2015), and during the night (Celiker Tosun 2015), with the PFMT group having higher number of micturition episodes during the day and during the night; and severity of impact on QoL, with the PFMT group reporting lower (Celiker Tosun 2015), and higher (Leong 2015), IIQ scores. For one trial, groups were different for the primary outcome (ICIQ-UI Short Form), with the PFMT group reporting lower severity of symptoms than the control group at baseline (Solberg 2016).

#### *Analysis by intention-to-treat, attrition and dropout*

Full intention-to-treat analysis requires that all participants are analysed in the group to which they were randomly assigned whether they adhered to treatment or not, crossed over to other treatments, or withdrew (Ferguson 2002). However, for the purpose of this review we have accepted the results as presented in the reports for those participants who provided outcome data at any time point, unless there was evidence of differential dropout from the groups. This was only the case in one trial, but we were unable to adjust the data (Sar 2009).

It was not clear if any other included study met the above definition/criteria for intention-to-treat, but four stated that the primary analysis was by intention-to-treat (Asklund 2017; Bidmead 2002; Burgio 1998; Sran 2016), while another stated that intention-to-treat analysis did not alter the findings of the primary analysis (Bø 1999). We have assumed that, in the absence of information to the contrary, all the trials analysed the participants in their assigned groups, with the exception of Sar 2009, as noted above.

Eight trials reported outcome data for all randomised participants (i.e. there appeared to be no dropouts) (Aksac 2003; Carneiro 2010; Ferreira 2014; Henalla 1989; Henalla 1990; Hofbauer 1990; Leong 2015; Miller 1998).

Five trials reported data only for those participants who reached outcome time points, but there was no evidence of differential dropout from the groups (Bertotto 2017; Diokno 2010; Kim 2011a; Kim 2011b; Pereira 2011).

One trial provided insufficient information to form an opinion on intention-to-treat analysis because the numbers at the outcome time points were not provided (Beuttenmuller 2010).

## Effects of interventions

See: **Summary of findings for the main comparison** Pelvic floor muscle training compared to control for stress urinary incontinence in women; **Summary of findings 2** Pelvic floor muscle training compared to control for urinary incontinence (all types) in women; **Summary of findings 3** Pelvic floor muscle training compared to control for urgency urinary incontinence in women

Thirty-one RCTs or quasi-randomised trials that compared PFMT (933 women) with no treatment, placebo, sham or other non-active control treatments (884 women) were eligible for inclusion in this review. Four trials did not contribute any data suitable for meta-analysis (Bidmead 2002; Ferreira 2014; Miller 1998; Wells 1999). In the 27 trials contributing data, the two comparison groups comprised 793 women in the PFMT group and 762 women in the comparison group.

In the following, as reflected by the subgroups in the forest plots, we presented the results for individual outcomes separately for the four incontinence subgroups: SUI (18 trials), UUI (one trial), MUI (one trial) and UI of all types (nine trials). Where available, we included subgroup data such as for participants with SUI from trials in the last category in the relevant diagnostic category.

Readers should note that when referring to the forest plots for some of the analyses, the right-hand side of the plot favours PFMT (Analysis 1.1; Analysis 1.2; Analysis 1.13; Analysis 1.14; Analysis 1.17; Analysis 1.18; Analysis 1.25; Analysis 1.26). For the remaining analyses, the left-hand side of the plot favours PFMT (Analysis 1.3; Analysis 1.4; Analysis 1.5; Analysis 1.6; Analysis 1.7; Analysis 1.8; Analysis 1.9; Analysis 1.10; Analysis 1.11 Analysis 1.12; Analysis 1.15; Analysis 1.16; Analysis 1.19; Analysis 1.20; Analysis 1.21; Analysis 1.22; Analysis 1.23 Analysis 1.24; Analysis 1.27; Analysis 1.28). This decision was made to keep interpretation of the forest plots clinically intuitive.

When a study measured an outcome but the data could not be included in the analysis for some reason, we noted this and briefly discussed the consistency with the usable data. Data in 'Other data' tables are only briefly discussed to give an indication of whether the findings were broadly consistent or not.

### Primary outcomes

#### *Participant-reported measures*

##### Symptomatic cure of urinary incontinence at the end of treatment

Six trials reported data on participant-reported cure of symptoms: four included women with SUI alone (Bø 1999; Hofbauer 1990; Kim 2007; Kim 2011a), and three included women with all types of incontinence (Burgio 1998; Kim 2011a; Kim 2011b). SUI subgroup data were available from Kim 2011a, which appears in both diagnostic categories. This outcome was not reported by the single trials for UUI or MUI. Although the CIs for most trials reporting data for SUI alone or all types of incontinence were wide, all trials found that women assigned PFMT were more likely to report that they were cured (Analysis 1.1).

##### Stress urinary incontinence alone

Women assigned PFMT were eight times more likely to report cure of symptoms than controls (46/82 (56%) with PFMT versus 5/83 (6%) with control; RR 8.38, 95% CI 3.68 to 19.07; 4 trials, 165 women; $I^2$ = 0%; high-quality evidence; Analysis 1.1.1).

##### Urinary incontinence (all types)

The three trials which included women with all types of UI showed a significant result favouring PFMT (50/144 (35%) with PFMT versus 9/146 (6%) with control; RR 5.34, 95% CI 2.78 to 10.26; 3 trials, 290 women; $I^2$ = 74%; moderate-quality evidence; Analysis 1.1.4). There was substantial statistical heterogeneity ($I^2$ = 74%), although there was agreement in the direction of effect in all three individually, favouring PFMT. The findings still favoured PFMT even when a random-effects model was used (RR 7.50, 95% CI 1.03 to 54.63). Visual inspection of the forest plot suggested a smaller effect size in Burgio 1998, while the effect size appeared similar in the two remaining trials (i.e. the magnitude of the benefit was uncertain but the analysis suggested consistency in the direction of effect). A possible explanation of this difference in treatment effect may come from the percentage of women with urgency symptoms, which was higher in Burgio 1998 than in the other two trials.

##### Symptomatic cure or improvement of urinary incontinence at the end of treatment

Five trials contributed outcome data for cure or improvement of symptoms: three included women with SUI alone (Asklund 2017; Bø 1999; Lagro-Janssen 1991a), and two included women with all types of UI (Burgio 1998; Diokno 2010). The single trials for UUI or MUI did not report this outcome. All five reported that PFMT was better than control interventions.

##### Stress urinary incontinence alone

Women assigned PFMT were six times more likely to report cure or improvement than women assigned control (88/119 (74%) with PFMT versus 14/123 (11%) with control; RR 6.33, 95% CI 3.88 to 10.33; 3 trials, 242 women; $I^2$ = 43%; moderate-quality evidence; Analysis 1.2.1).

##### Urinary incontinence (all types)

Women assigned PFMT were twice as likely to report cure or improvement than women assigned control (58/86 (67%) with PFMT versus 23/80 (29%) with control; RR 2.39, 95% CI 1.64 to 3.47; 2 trials, 166 women; $I^2$ = 0%; moderate-quality evidence; Analysis 1.2.4).

Wells 1999 reported information on cure or improvement with no difference between treatment groups, but the data were not suitable for meta-analysis (no raw data available). Leong 2015 reported on participant perception of improvement on a visual analogue scale (0 to 10: higher scores = better outcome): this favoured PFMT (MD 7.30, 95% CI 6.84 to 7.76; 55 women; Analysis 1.18.4).

##### Symptom- and condition-specific quality of life measures

Grade A UI symptoms and QoL measures are presented in the forest plots. Eight out of nine different measures of QoL specific to the effect of UI were in favour of PFMT in women with SUI, MUI and all types of UI (Analysis 1.3; Analysis 1.4; Analysis 1.5; Analysis 1.7; Analysis 1.8; Analysis 1.9; Analysis 1.10; Analysis 1.11; Analysis 1.12). Evidence of higher QoL following PFMT was not evident in the

three trials that reported the King's Health Questionnaire general health score in women with SUI, but this may be because measures of general health are less sensitive to changes in continence (Beuttenmuller 2010; Carneiro 2010; Pereira 2011; Analysis 1.6).

In the King's Health Questionnaire, which measures the impact of incontinence after treatment, there was considerable statistical heterogeneity ($I^2$ = 76%). When a random-effects model was used there was no evidence of a difference between treatment groups, although all trials had the same direction of effect and their CIs included clinically important differences (Analysis 1.4). Visual inspection of the forest plot suggested a larger effect size in Pereira 2011, while the effect size appeared similar in the two remaining trials. A possible explanation of this difference in treatment effect may come from the intensity of the PFMT programme, which was higher in Pereira 2011 than in the two others. In the IIQ short form there was considerable statistical heterogeneity ($I^2$ = 99%) and, although all trials were on the same side of the forest plot, we decided not to combine them in a meta-analysis considering that the baseline comparability for the IIQ short form score was at high risk of bias in both trials (Celiker Tosun 2015; Leong 2015; Analysis 1.10.4).

Three trials reported lower grade measures of UI symptoms and QoL (Bø 1999; Diokno 2010; Firra 2013). Their effects are given in detail in Appendix 4.

## Secondary outcomes

### *Participant-reported measures*

#### Longer-term symptomatic cure and improvement

There was limited information from low- to moderate-quality trials indicating that the benefit of PFMT seemed to persist after treatment stopped for up to one year for cure in women with all types of UI (23/59 (38.9%) with PFMT versus 1/61 (1.6%) with control; RR 23.78, 95% CI 3.32 to 170.49; Analysis 1.13; Kim 2011b), and for cure and improvement in women with SUI only (14/26 (53.8%) with PFMT versus 0/25 (0%) with control; RR 27.93, 95% CI 1.75 to 444.45; Analysis 1.14; Henalla 1989). The CIs in both trials were wide and hence these results need further confirmation.

#### Longer-term symptom- and condition-specific quality of life measures

One trial indicated that the benefit of PFMT seemed to persist (after treatment stopped) for up to one year in women with UI (all types) in regards to symptoms (UDI long form) (69.15 (SD 38.14) with PFMT versus 107.73 (SD 61.72) with control; MD −38.58, 95% CI −67.61 to −9.55; Analysis 1.15; Sran 2016), and UI-specific QoL measures (IIQ long form) (20.44 (SD 30.71) with PFMT versus 62.35 (SD 98.54) with control; MD −41.91, 95% CI −83.20 to −0.62; Analysis 1.16; Sran 2016). One trial published one-year and two-year follow-up reports with data on symptoms (ICIQ-UI Short Form) and QoL (ICIQ-LUTSqol) but only for the PFMT group (Asklund 2017).

#### Satisfaction

Three trials measured participant satisfaction with treatment for SUI (Bø 1999; Castro 2008), or for women with UI (all types) (Burgio 1998; Analysis 1.17). In trials which included women with SUI alone, women in the PFMT group were five times more likely to be satisfied with the intervention than women in the control group (36/51 (70.6%) with PFMT versus 7/54 (12.9%) with control; RR 5.32, 95% CI 2.63 to 10.74; Analysis 1.17.1). There was substantial statistical heterogeneity ($I^2$ = 74%) but the findings still favoured PFMT when a random-effects model was used (RR 5.54, 95% CI 1.15 to 25.63). In the one trial with women with all types of UI, women assigned to PFMT were three times more likely to be satisfied with the intervention than women assigned to control (45/58 (78%) with PFMT versus 14/50 (28%) with control; RR 2.77, 95% CI 1.74 to 4.41; Analysis 1.17.4).

#### Need for further treatment

Two trials reported that more women needed further treatment in the control groups; one trial in women with SUI (RR 0.17, 95% CI 0.07 to 0.42; Bø 1999), and one in women with UI of all types (RR 0.19, 95% CI 0.10 to 0.36; Burgio 1998) (Analysis 1.19).

#### Perception of improvement

One trial measured perception of improvement in women with UI (all types) with women assigned to PFMT reporting perception of improvement seven times higher than women assigned to control (8.7 (1) with PFMT versus 1.4 (0.7) with control; MD 7.30, 95% CI 6.84 to 7.76; Analysis 1.18.4; Leong 2015).

#### Self-efficacy

One trial measured self-efficacy (Sran 2016). In women with SUI, the PFMT group presented higher self-efficacy immediately after treatment and at one year compared to the control group (after treatment: MD 0.19, 95% CI 0.06 to 0.32; at one year: MD 0.15, 95% CI 0.02 to 0.28).

### *Participant-reported quantification of symptoms*

#### Number of leakage episodes (in 24 hours)

Twelve trials contributed to the forest plot: seven for the SUI subgroup (Asklund 2017; Bø 1999; Burns 1993; Castro 2008; Firra 2013; Lagro-Janssen 1991a; McLean 2013); one for the UUI subgroup (Firra 2013); and four for the UI of all types subgroup (Burgio 1998; Celiker Tosun 2015; Leong 2015; Sran 2016). While the overall effect for the SUI subgroup showed a clinically important difference favouring PFMT, visual inspection of the forest plot suggested the effect size might be greater in the trial by Lagro-Janssen 1991a, while the effect sizes appeared similar in the six remaining trials (Analysis 1.20.1). It was not clear why the data from Lagro-Janssen 1991a were different from the other six trials in women with SUI, or the trials overall. A possible explanation of the greater treatment effect might be an inadequate concealment of the randomisation process. The point estimates in the other six trials were similar and all provided evidence of a clinically important reduction in leakage episodes following PFMT, except for one (Firra 2013). There was substantial heterogeneity ($I^2$ = 73%), but the findings still favoured PFMT with the random-effects model used. Women with SUI assigned PFMT experienced about one leakage episode less in 24 hours compared to women assigned control (MD −1.23, 95% CI −1.78 to −0.68; 7 trials, 432 women; moderate-quality evidence; Analysis 1.20.1). For women with UUI, one trial favoured PFMT over control (0.77 (SD 0.97) with PFMT versus 2.6 (SD 0.33) with control; MD −1.83, 95% CI −2.65 to −1.01; 12 women; low-quality evidence; Analysis 1.20.2; Firra 2013). Similarly, according to four trials, women with UI of all types assigned to PFMT experienced about one less leakage episode in 24 hours compared to women assigned to control (MD −1.00, 95% CI −1.37 to −0.64; 349 women; moderate-quality evidence; Analysis 1.20.4).

#### Number of micturitions during the day (frequency)

Four trials reported data on number of micturitions during the day (frequency) (Celiker Tosun 2015; Diokno 2010; Firra 2013; Yoon 2003). One trial reported on frequency for women with SUI (MD –0.60, 95% CI –3.09 to 1.89; Analysis 1.21.1; Firra 2013), and UUI (MD –0.24, 95% CI –3.43 to 2.95; Analysis 1.21.2; Firra 2013) separately, without a statistically significant difference between groups. In three trials, women assigned PFMT with all types of UI reported about two fewer micturitions per day than women assigned control (MD –2.32, 95% CI –3.21 to –1.43; Analysis 1.21.4; Celiker Tosun 2015; Diokno 2010; Yoon 2003).

#### Number of micturitions during the night (nocturia)

For the number of micturitions during the night, in three trials there was no evidence of a difference between the PFMT and control groups and the CI was wide (MD –0.03, 95% CI –0.46 to 0.40; Analysis 1.22.4; Celiker Tosun 2015; Diokno 2010; Yoon 2003).

Of interest might be the fact that five trials reported other leakage outcomes rather than the number of leakages or number of micturitions (one leakage index, two leakage scores, number of urgency episodes and urine leakage at one year) (Bø 1999; Celiker Tosun 2015; Kim 2011b; Sran 2016; Yoon 2003). These are reported in detail in Appendix 5.

### *Clinicians' measures*

#### Pad and paper towel tests short (up to one hour) or long (24 hours) urine loss at the end of treatment

#### Up to one hour

Four trials reported urine loss on pad tests in women with SUI (Bø 1999; Castro 2008; McLean 2013; Pereira 2011), and two in women with UI (all types) (Celiker Tosun 2015; Yoon 2003). Women with SUI in the PFMT groups lost significantly less urine on the up to one hour pad tests (MD –4.22 g, 95% CI –6.56 to –1.88; 185 women; moderate-quality evidence; Analysis 1.23.1). There was statistical considerable heterogeneity ($I^2$ = 78%), but the finding still favoured PFMT if a random-effect model was used (MD –9.71 g, 95% CI –18.92 to –0.50). Visual inspection of the forest plot suggested that the effect size might be greater in one trial (Bø 1999). One possible explanation for this difference would be the variation among test protocols. Bø 1999 used a 60-second pad test in which the participants were required to run on the spot for 30 seconds and then jump with legs in subsequent adduction and abduction (jumping jacks) at a rate of 132 beats/minute. For the other three trials, the pad test protocol appeared to be more similar, varying from 30 minutes to one hour, in which the participants were asked to do different circuits, including walking, running, jumping and coughing (Castro 2008; McLean 2013; Pereira 2011). For women with unspecified UI, the PFMT groups also reported less urine loss than the control groups (MD –3.72 g, 95% CI –5.46 to –1.98; 2 trials, 146 women; moderate-quality evidence; Analysis 1.23.4).

#### Over 24 hours

One trial reported urine loss on a 24-hour pad test with women with SUI (Bø 1999), and two trials with women with UI of all types (Diokno 2010; Sran 2016). There was no difference between PFMT and control on the 24-hour test for the two subgroups (SUI or all types of UI) (Analysis 1.24). One trial reported urine loss on a 24-hour pad test after one-year follow-up, favouring PFMT (MD –30.50, 95% CI –55.97 to –5.03; Sran 2016; Appendix 8).

#### Number cured or improved based on pad weights in short clinic-based pad test at the end of treatment

When urine leakage was objectively assessed based on the number of women who had dry pads (short pad test), women with SUI were more likely to be cured in the PFMT groups (number cured: 38/71 (53.5%) with PFMT versus 4/64 (6.3%) with control; RR 7.50, 95% CI 2.89 to 19.47; Analysis 1.25.1), and similarly for cure or improvement (41/54 (75.9%) with PFMT versus 2/42 (4.8%) with control; RR 8.22, 95% CI 3.17 to 21.28; Analysis 1.26.1).

Four trials reported pad or paper towel tests in other ways or reported data where the MD was not estimable (Aksac 2003; Bidmead 2002; Diokno 2010; Miller 1998). These data are given in detail in Appendix 8. The data were generally in agreement with the findings above.

### *Quality of life (not condition specific)*

#### General health status measures

Validated measures were used to assess generic QoL (Bø 1999), psychological distress (Burgio 1998), and self-esteem (Kargar Jahromi 2013). There was no statistically significant difference between PFMT and control groups in either women with SUI or women with all types of UI for generic QoL or psychological distress. Women with SUI in the PFMT group presented higher self-esteem after treatment compared to the control group (MD 5.28, 95% CI 2.71 to 7.85; Appendix 6).

#### Psychosocial outcome measures

None of the trials reported psychological outcome measures.

#### Sexual function or problems

One trial with women with SUI suggested that PFMT improved the women's sex life both generally and in terms of reduction of urine leakage during intercourse (Analysis 1.27; Analysis 1.28; Bø 1999).

### *Adverse effects*

Seven trials specifically mentioned adverse events, and five did not report any in the PFMT group (Bø 1999; Burgio 1998; Castro 2008; Leong 2015, Sran 2016). Two trials reported adverse events with PFMT (Lagro-Janssen 1991a; Solberg 2016). These were: worsening of incontinence symptoms after the first two treatments that disappeared as treatment continued (one woman; Solberg 2016), or pain (one woman); uncomfortable feeling during exercise (three women); and "not wanting to be continuously bothered with the problem" (two women; Lagro-Janssen 1991a).

### *Measures of likely moderator variables*

#### Measures of pelvic floor muscle function

Fifteen trials reported measures of pelvic floor muscle function (Aksac 2003; Bertotto 2017; Beuttenmuller 2010; Bø 1999; Burns 1993; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Diokno 2010; Firra 2013; McLean 2013; Miller 1998; Pereira 2011; Wells 1999; Yoon 2003).

- Three trials used ultrasound to measure morphological changes in pelvic floor muscles after treatment (Carneiro 2010; Celiker Tosun 2015; McLean 2013).
- Seven trials used vaginal squeeze pressure to measure functional changes in pelvic floor muscles (Aksac 2003;

Beuttenmuller 2010; Bø 1999; Celiker Tosun 2015; Firra 2013; Pereira 2011; Yoon 2003).

- Nine trials used vaginal digital assessment to measure functional changes in pelvic floor muscles (Aksac 2003; Beuttenmuller 2010; Carneiro 2010; Castro 2008; Celiker Tosun 2015; Diokno 2010; Miller 1998; Pereira 2011; Wells 1999).
- Four trials used electromyography (EMG) measures of pelvic floor muscle function (Bertotto 2017; Burns 1993; Carneiro 2010; Wells 1999).

Of the 15 trials, three did not report the data in such a way that it was possible to calculate the MD in PFM morphometry as measured by ultrasound, vaginal squeeze pressure, EMG activity or digital palpation score (Aksac 2003; McLean 2013; Wells 1999). Overall, there were no consistent patterns in measures of pelvic floor muscle function. Details are given in Appendix 9.

#### Measures of adherence

#### From diaries

Leong 2015 reported the highest rate of adherence to PFMT (99.4%) and Bø 1999 reported the second highest rate of adherence to PFMT (93%), both using exercise and training diaries. Further, Castro 2008 reported the third highest rate of adherence to PFMT (92%), using a training diary. Sran 2016 reported a high rate of adherence to PFMT using an exercise diary in the PFMT group with 33% (8/24) of participants completing 100% of the 12-week home exercises and 33% (8/24) completing 70% to 99% of the home exercises programme. At one year, 78% (18/23) of the physiotherapy participants continued to do the PFM exercises using an exercise diary. Bidmead 2002 found that 75% of women assigned to PFMT had excellent (daily) or good (training more than three times per week) adherence to exercise on using exercise and training diaries. Women in the study by Lagro-Janssen 1991a rated their adherence as excellent or good (62%), reasonable (20%), or poor or none (18%). Kim 2007 reported adherence to home PFMT only in the follow-up period (after the intervention to the follow-up assessment) using exercise and training diaries, with 30% of women doing their pelvic floor muscle exercises every day, two to three times per week in 45.5%, and once or less per week in 24.2%. In Kim 2011b, the same research group reported adherence using exercise and training diaries for home PFMT in the follow-up period, again with 35.7% of women doing their pelvic floor muscle exercises every day, two to three times per week in 42.9% and once or less per week in 21.4%. Wells 1999 reported a greater exercise frequency in the treatment group at the beginning of the trial, although no raw data were available to support this finding. Asklund 2017 was able to compute the completion of each PFMT exercise and save in a statistics table in the application, reporting that 41.0% (25/61) had performed PFMT daily.

#### From attendance at appointments

Five trials attempted to measure attendance at exercise sessions (Burns 1993; Castro 2008; Kim 2007; Leong 2015; Sran 2016). Four trials reported very good to excellent attendance rates at clinic appointments (70%, Kim 2007; 92%, Castro 2008; 92%, Sran 2016; 98%, Leong 2015), and the fifth did not present any data (Burns 1993).

#### Methods to increase adherence

Eight trials used adherence strategies to encourage participants to do their PFMT exercises (Asklund 2017; Bø 1999; Burns 1993; Celiker Tosun 2015; Diokno 2010; Kim 2007; Sar 2009; Solberg 2016). Sar 2009 used a telephone call to encourage participants and answer questions. Diokno 2010 used as reinforcement a two- to four-week follow-up which consisted of vaginal examination, measurement of pelvic floor muscle strength and a test measuring participants' ability to correctly perform the verbally instructed exercise programme. Burns 1993 used weekly and three- and six-month telephone reminders for treatment appointments and weekly exercise reminder cards were mailed between visits. Bø 1999 used audiotape with verbal guidance for home training. Kim 2007 used a pamphlet illustrating the pelvic floor muscles and strengthening exercises. Two trials reported that participants were required to keep a training diary to maintain their motivation (Celiker Tosun 2015; Solberg 2016). There were no data on adherence. Asklund 2017 included adherence strategies in their application to promote adherence (reminder setting that was used by 83.6% (51/61) and the statistics function used by 86.9% (53/61)).

#### GRADE quality of evidence

'Summary of findings' tables were prepared separately for women with SUI at baseline (Summary of findings for the main comparison), for women with all types of UI (SUI, UUI, MUI) (Summary of findings 2), and for women with UUI (Summary of findings 3). The findings of the review were supported in the tables but in all cases except one the quality of the evidence was downgraded. The exception was 'Participant perceived cure – stress urinary incontinence,' which was rated as high-quality. This suggested that SUI was eight times more likely to be cured in this subgroup (RR 8.38, 95% CI 3.68 to 19.07; Analysis 1.1.1), which is a much higher estimate of success than suggested in the other subgroups or other outcomes. However, although the CI was wide and was derived from two small- and two moderate-size trials, we can be confident that PFMT does improve outcomes.

## DISCUSSION

This review is one of a series of reviews of PFMT for UI in women and should be viewed in that context. Other reviews considered whether:

- one type of PFMT was better than another (Hay-Smith 2011), or whether feedback or biofeedback had a role to play (Herderschee 2011);
- PFMT was better than other treatments (e.g. other physical therapies, medication and surgery) (Lins 2014); and
- if the addition of PFMT to other therapies added benefit (Ayeleke 2015).

A separate review considered the role of PFMT in the treatment and prevention of urinary and faecal incontinence related to childbirth (Woodley 2017).

### Summary of main results

#### Is pelvic floor muscle training better than no treatment, placebo or control treatments?

Of the 31 trials that addressed this question, 27 reported data suitable for analysis for the outcomes of interest.

*Symptomatic cure of urinary incontinence at the end of treatment:* compared with no treatment or to inactive control treatments, women with SUI who were in the PFMT groups were eight times

more likely to report cure (56% with PFMT versus 6% with control; RR 8.38, 95% CI 3.68 to 19.07; 4 trials, 165 women; high-quality evidence; Analysis 1.1.1). For women with any type of UI, PFMT groups were five times more likely to report cure (35% with PFMT versus 6% with control; RR 5.34, 95% CI 2.78 to 10.26; 3 trials; 290 women; moderate-quality evidence; Analysis 1.1.4), although with substantial statistical heterogeneity ($I^2$ = 74%). However, using a random-effect model, the results would still favour PFMT. Visual inspection of the forest plot suggested a smaller effect size in one trial (RR 2.34, 95% CI 1.11 to 4.94; Burgio 1998), where the urgency component of UI was more prevalent than in the two other trials (RR 11.48, 95% CI 0.67 to 196.07; Kim 2011a; RR 26.88, 95% CI 3.77 to 191.79; Kim 2011b).

*Symptomatic cure or improvement of urinary incontinence at the end of treatment:* compared with no treatment or to inactive control treatments, women with SUI who were in the PFMT groups were six times more likely to report cure or improvement (74% with PFMT versus 11% with control; RR 6.33, 95% CI 3.88 to 10.33; 3 trials, 242 women; moderate-quality evidence; Analysis 1.2.1). For women with any type of UI, PFMT groups were two times more likely to report cure or improvement than the women in the control groups (67% with PFMT versus 29% with control; RR 2.39, 95% CI 1.64 to 3.47; 2 trials; 166 women; moderate-quality evidence; Analysis 1.2.4).

Where reported, QoL due to incontinence was also improved by the active PFMT intervention in women with SUI, women with MUI and women with all types of UI (moderate-quality evidence). Women were also more satisfied with the active treatment, while women in the control groups were more likely to seek further treatment.

*Leakage episodes in 24 hours*: PFMT reduced leakage episodes by one in women with SUI (MD 1.23 lower, 95% CI 1.78 lower to 0.68 lower; 7 trials, 432 women; moderate-quality evidence), and in women with all types of UI (MD 1.00 lower, 95% CI 1.37 lower to 0.64 lower; 4 trials, 349 women; moderate-quality evidence). PFMT may have reduced leakage episodes by 1.8 in women with UUI (MD 1.83 lower, 95% CI 2.65 lower to 1.01 lower; 1 trial, 12 women; low-quality evidence).

Women assigned PFMT also lost smaller amounts on short clinic-based pad tests, emptied their bladders less often during the day and their sexual outcomes were better. Adverse events were rare and, in the two trials that did report any, they were minor. However, there was no evidence of a change in QoL measures, perhaps because measures of general health are less sensitive to changes in continence status or because there was insufficient evidence.

The improvement in pelvic floor muscle function as the mechanism by which UI improved was supported by many trials. Attendance at treatment sessions was generally good, and women were also motivated to practice their pelvic floor exercises during the intervention period. However, only three trials presented information about persistence of benefit in the long-term, and the need for further treatment such as incontinence surgery or drugs was scant.

The findings of the review were largely supported in the 'Summary of findings' tables, but most of the evidence was downgraded to moderate on methodological grounds (Summary of findings for the main comparison; Summary of findings 2; Summary of findings 3).

For a future update, we intend to revisit the list of outcomes in accordance with current Cochrane guidance. First, we will pull all UI-specific symptoms and QoL data from different grade A questionnaires into two outcomes, which will be presented in the forest plots and not only as a narrative report in the 'Summary of findings' tables. This will give a better overview of combined UI-specific symptoms and QoL outcomes, and will reduce the number of forest plots and appendices.

Further, although we think adherence is a very important measure and we encourage authors to report adherence, we intend to remove it from the outcomes of PFMT list. Finally, non-specific QoL outcomes, adherence outcomes and PFM function assessments, although very important measures, will be removed from the list of outcomes as they are not directly related to PFMT for the reduction of UI.

## Overall completeness and applicability of evidence

### Types of incontinence

Although we prespecified four clinical subgroups for baseline type of UI (SUI, UUI, MUI and UI of all types) in the analysis, most of the trials reported data on two of them (SUI, UI of all types). Only two trials investigated the effect of PFMT versus control in the two remaining subgroups, one in women with UUI only and another in women with MUI only.

Further, participants were selected for the trials solely on the basis of the type of incontinence, diagnosed according to signs, symptoms or urodynamics. Theoretically, women with ruptured ligaments or fascia, partial or complete avulsion of the PFM, or even severe peripheral nerve damage may have responded differently to PFMT than women without such major anatomical defects, which may affect the estimate of treatment effect. Using new imaging techniques may improve researchers' ability to give a more specific diagnosis and use a more homogenous sample of participants, or present their data according to women who did and did not have such defects (Dumoulin 2011).

### Variation in interventions

There was large variation in the PFMT programmes, as reported in Appendix 3. Further, the exercise regimen in both the clinic-based and home PFMT programmes was often incompletely reported. It was difficult to make judgements about the similarities and differences between the training programmes, and hence their potential relative effectiveness. Including trials with a suboptimal exercise 'dose' could adversely affect the estimate of differences in treatment effect. Although assessment of the interactions between the quality of the exercise programmes and their effects has been recommended (Herbert 2005), it was not possible to explore this aspect in this review. Nevertheless, the more recent trials reported PFMT exercise regimens that were more in line with the literature on skeletal muscle training theory and pelvic floor muscle dysfunction, with supervised progressive training protocols.

### Outcomes

Some important secondary outcomes were either missing or were rarely reported. Only three trials reported medium-term follow-up (less than one year), all of which favoured the active PFMT but with very wide CIs (Analysis 1.13; Analysis 1.14). Although one new trial reported data for one- and two-year follow-up, it was only given for one group (the PFMT group). Ethical reasons may

prevent maintenance of the control group in a no treatment state and may explain this absence of longer-term follow-up. Arguably, the need for further treatment (e.g. the use of another conservative intervention, pessary, surgery or a drug) would provide a robust and objective measure of the ultimate success of treatment and should be reported in future studies. Unfortunately, this was not reported in any of the trials.

Treatment adherence (e.g. performance of pelvic floor muscles exercise) was reported only in the short-term (during the intervention) and in some trials not in the control groups, so it could not be compared between the groups. Therefore, it was not possible to assess the interactions between the effect size and the adherence to treatment.

## Quality of the evidence

### Trial quality and reporting

A total of 31 mostly small-to-moderate trials were included in the review for the SUI and all UI population subgroups (1793 women). One trial contributed to the UUI only subgroup (12 women) and another to the MUI only subgroup (12 women).

The major limitation in the reporting of included trials was the absence of a clear description of the PFMT programmes. Another problem was the absence of long-term follow-up or need for further treatment.

The results were consistent for most of the outcomes, favouring PFMT over control. The only outcome that was consistently not different between the experimental and control conditions was generic QoL, but these measures may not be sensitive enough to pick up changes due to improvement in UI.

### 'Summary of findings' tables and the GRADE approach

The main reasons for downgrading the quality of the evidence in the 'Summary of findings' tables were:

- random sequence generation and allocation concealment was high or unclear risk in some trials;
- baseline comparability was high or unclear risk in some trials;
- results were imprecise (heterogeneity due to variation in results, although these were generally in favour of PFMT); and
- results were imprecise (for being self-reported measures).

## Potential biases in the review process

Of the 31 included trials, nine were at high risk of bias:

- Diokno 2010 for its differences in baseline comparability (especially with regard to age; those in the treatment group were older);
- Ferreira 2014 for not reporting all expected outcomes in the results section;
- Firra 2013 for its lack of blinding of outcome assessment, and also for its differences in baseline comparability (groups were different only in the UUI subgroup, with those in the control group presenting lower frequency in the three-day diary);
- Henalla 1990 for not reporting all expected outcomes in the results section;
- Lagro-Janssen 1991a for its lack of genuine randomisation and inadequate allocation concealment;
- Leong 2015 for its lack of blinding of outcome assessment, and for its differences in baseline comparability (those in the control group were older, and presented lower IIQ short form scores – less impacted);
- Pereira 2011 for its lack of blinding of outcome assessment;
- Sar 2009 for its management of attrition; and
- Solberg 2016 for its differences in baseline comparability (with those in the PFMT group presenting lower ICIQ-UI Short Form score), and for attrition bias, considering that almost half of the sample were lost to follow-up, with no intention-to-treat analysis.

Because of the nature of the intervention, which is a complex interaction between the therapist and the patient, it was not possible to blind either party and, therefore, we rated all trials as unclear for this domain. It was also difficult to assess incomplete outcome data because most of the trials did not publish their protocols. We rated trials at low risk of bias if the expected outcomes were presented in the results section. Only Asklund 2017, Bertotto 2017, Leong 2015, Pereira 2011, and Sran 2016 published a protocol. Many trials were not able, and did not intend to, report long-term follow-up. In nine trials, the authors stated that the untreated groups would receive treatment after the end of the trial (Asklund 2017; Bertotto 2017; Bidmead 2002; Burns 1993; Celiker Tosun 2015; Kim 2007; Miller 1998; Pereira 2011; Solberg 2016).

We combined data from a diverse set of studies. This may inevitably impact on the applicability of our findings to practice. Finally, we were not able to use Egger's test to assess publication bias in the current version of the review. However, we plan to include it in future updates when there are more than 10 studies per subgroup analysis.

### Sources of heterogeneity

#### *Variability of pelvic floor muscle training regimens*

PFMT programmes varied considerably in their content and duration across trials (Appendix 3). Furthermore, the exercise regimen in both the clinic-based and home PFMT programmes were often incompletely reported. Therefore, it was difficult to make judgements about the similarities and differences between interventions, and hence their potential relative effectiveness. However, although the training protocol was poorly described in many of the older trials (6/31 trials), the more recent trials presented reports of PFMT exercise regimens that were more in line with the literature on skeletal muscle training theory and pelvic floor muscle dysfunction, with supervised (26/31 trials) and progressive exercises (14/31 trials). The duration of the programmes varied between one week and six months, though most of the trials had programmes ranging between eight and 12 weeks (20/31 trials). Consequently, it was difficult to evaluate the potential physiological effect of the exercise programmes. Including trials with a suboptimal exercise regimen alongside those with a sufficient regimen could adversely influence the pooled estimate of PFMT effect. Alongside the physiological effect of the exercise, support for the behavioural aspects of exercise is also required. Behavioural support is commonly provided through supervision of exercise, and the extent of this varied markedly between trials. Most of the trials provided at least weekly supervised PFMT classes (21/31). The least supervision was either having no face-to-face interaction with a health professional (one trial) or one session to confirm a correct PFM contraction prior to

training (three trials). The most was a mean of 72 classes (three classes per week) over six months (one trial).

Assessment of the interaction between quality and the effect of the intervention has been recommended but there were too few trials to conduct a formal sensitivity analysis by intervention quality (Herbert 2005). As such, it was not possible to explore this aspect in this review. Rather than excluding or including trials on the basis of sufficiency of PFMT, or the likelihood that a clear-cut comparison between PFMT and the control condition had been made, the preferred approach would have been to conduct a sensitivity analysis on the basis of PFMT programme characteristics or amount of clinical difference between the PFMT and control interventions. However, more trials would be needed in each of the comparisons in the review before this was possible.

#### *Variability of control conditions*

Control conditions were also highly variable. They included no treatment (19 trials), placebo drug (one trial), sham electrical stimulation (one trial), and a variety of inactive intervention strategies, including educational pamphlets, general education classes or advice on lifestyle alterations. However, it was often unclear whether the control group was advised on PFMT, or if they were doing home PFMT exercises (see the Characteristics of included studies).

### Agreements and disagreements with other studies or reviews

The findings of this update are consistent with the previous version of this Cochrane Review (Dumoulin 2014), and a Heath Technology Assessment monograph which investigated all conservative methods of managing SUI (Imamura 2010).

#### Brief Economic Commentary

To supplement the main systematic review of the efficacy of pelvic floor muscle exercises in the treatment of UI, we identified economic evaluations comparing the intervention to a placebo or a sham control. One cost utility analysis was identified (Sjostrom 2017, which was a further report of Asklund 2017; see the Characteristics of included studies and Appendix 3 for details of the intervention and control). The analysis claimed to adopt a societal perspective over a one-year time horizon. The costs included administration costs for running the app and lost earnings for the women while doing the exercises. The development costs for the application were excluded. The outcomes were expressed as ICIQ-UI Short Form and ICIQ-LUTSqol scores. These scores were mapped to utility values using a preference-based index. The authors of the evaluation reported that the application providing instructions for PFMT was a cost-effective first-line treatment alternative.

We did not critically appraise the economic evaluation and we do not attempt to draw any firm or general conclusions regarding the relative costs or efficiency of the PFMT interventions. However, this evaluation does provide some evidence that application-based PFMT is a promising strategy for the management of UI. End users of this review will need to assess the extent to which methods and results of the economic evaluation may be applicable or transferable to their own setting.

#### Considerations for future research

The outcomes of incontinence research would be much more useful if trialists selected a primary outcome measure that mattered to women, chose secondary measures to cover a range of important domains, and opted for standardised tools with established validity, reliability and responsiveness to measure outcomes. One domain that requires particular attention in future is socioeconomics, as it has been poorly addressed to date. Three trials included in the review asked women if they wanted further treatment or were satisfied with the treatment outcome, or both. Questions such as these have potential merit, but asking women if they are cured or improved with treatment may not differentiate those who are improved and do not want any further intervention from those who are improved but not sufficiently so to be satisfied with the treatment outcome. Although better in the most recent trials, there is also scope for the use of validated questionnaires that evaluate bother or distress associated with symptoms (e.g. the Urogenital Distress Inventory, ICIQ-LUTSqol) instead of general health questionnaires that are less sensitive to changes in continence.

Duration of follow-up beyond the end of supervised treatment needs attention. As the aim of treatment is long-term continence, it would be appropriate if the outcome was measured at least one year after the end of treatment. As PFMT generally precedes other more invasive treatment options, such as surgery, the proportion of women satisfied with the outcome of PFMT (and for how long they remain so) would provide essential information for women, clinicians and service planners.

The reporting of methods and data could be improved. Some included trials collected data for outcomes of interest but did not report it in a useful manner (e.g. point estimates without a measure of dispersion). It was also difficult to assess one of the primary ways to minimise risk of bias for allocation concealment, because the methods of randomisation were often poorly described. Trialists are referred to the CONSORT and revised CONSORT statements for appropriate standards of trial reporting (Boutron 2008; Moher 2001).

In essence, there is a need for at least one large, pragmatic, well-conducted and explicitly reported trial comparing PFMT with control to investigate the longer-term (more than one year), clinical effectiveness and cost-effectiveness of PFMT. An important outcome measure should be added to cure and improvement of incontinence: the need to use extra interventions (such as pessaries, drugs or surgery) after the end of the PFMT intervention.

Such a trial could recruit separate groups of women with symptoms of SUI, UUI or mixed UI based on clinical history and physical examination, with a sample size based on a clinically important difference in self-reported UI and condition-specific QoL outcomes, and sufficient for subgroup analysis on the basis of type of UI. Stratification or minimisation procedures could be used to ensure an even distribution of women with different types of UI across both arms of the trial.

One arm of the study would comprise of a supervised PFMT programme based on sound exercise science with confirmation of a correct VPFMC, and incorporate appropriate supervision and adherence measures to promote maintenance of knowledge acquisition, behaviour skills and motivation (Dumoulin 2011). The

choice of programme would have to be set against the resource implications of intensively supervised individual programmes and the opportunity cost this represents. Careful clinical judgement is needed about what sort of programme could actually be applied in everyday practice and in different countries with their different healthcare delivery systems. The other arm of the trial would be a control treatment, for example an explanation of the anatomy and physiology of the bladder and pelvic floor, or advice on good bladder and lifestyle habits, with the same explanation and advice given in both arms. Such a trial would require substantial funding and multiple recruitment centres. A formal economic analysis, and process evaluation (e.g. to check intervention fidelity), would also be an important part of such a trial.

## AUTHORS' CONCLUSIONS

### Implications for practice

Based on the data available, pelvic floor muscle training (PFMT) is better than no treatment, placebo or inactive control treatments for women with stress urinary incontinence (SUI) or urinary incontinence (UI) (all types). We can be confident that PFMT can cure or improve symptoms of SUI and all other types of UI.

PFMT may reduce the number of leakage episodes in women with SUI and in women with all types of UI.

There was new information from two small trials about women with urgency urinary incontinence (UUI) alone or mixed urinary incontinence (MUI) also supporting PFMT. From this new addition, there is now low-quality evidence that PFMT reduces leakage episodes in women with UUI. For women with MUI treated with PFMT, there is now one report of better quality of life.

Overall, women with SUI or all types of UI treated with PFMT were more likely to report better quality of life and have less urine leakage on short clinic-based pad tests than controls. Women were also more satisfied with the active treatment, and their sexual outcomes were better. Finally, the findings of the review suggest that PFMT could be included in first-line conservative management programmes for women with UI.

The limited nature of follow-up beyond the end of treatment in the majority of the trials means that the long-term outcomes of the use of PFMT remain uncertain. At this time, we are only starting to gather data on whether PFMT is cost-effective in the long-term.

### Implications for research

There is a need for a pragmatic, well-conducted and explicitly reported trial comparing PFMT with control to investigate the longer-term clinical effectiveness and cost-effectiveness of PFMT for women with symptoms of SUI, UUI or MUI. Although the quality of recent trials has improved (choice of outcome, duration of follow-up, reporting method and data), most of the data in this review comes from small- to moderate-size trials of moderate methodological quality. In planning future research, trialists are encouraged to consider the following.

- The choice of primary outcomes important to women (urinary outcomes and quality of life), the size of a clinically important effect, and subsequent estimation of sample size.
- Choice and reporting of PFMT exercise programmes, including details of the number of voluntary pelvic floor muscle contraction per set, duration of hold, duration of rest, number of sets per day, body position, types of contractions and other recommended exercises.
- The reporting on adherence outcomes and adherence strategies, including practice of pelvic floor muscle exercises in both the intervention and control groups.
- The need for further treatment, such as with pessaries, surgery or drugs.
- The choice and reporting of secondary outcome measures, such as sexual function.
- The duration of follow-up, especially long-term follow-up.
- The reporting of formal economic analysis, such as cost-effectiveness and cost utility.

## ACKNOWLEDGEMENTS

We are grateful to Sue Bennett, Fiona Beyer, Gill Gyte, Emily Karahalios, Juliana Esther Martin-Lopez and Heather Welford for valuable comments on drafts of this version of the review. We also appreciate the work of Ashleigh Kernohan in conducting the Brief Economic Commentary; and Sheila Wallace, Lindsey Elstub and Luke Vale to this version.

We would also like to acknowledge the considerable contribution made by Cathryn Glazener, Muhammad Imran Omar, Suzanne MacDonald and Sheila Wallace to the previous version of this review in 2014.

We would also like to acknowledge the contribution of Kari Bø, Bary Berghmans, Erik Hendriks, Rob de Bie and Ernst van Waalwijk van Doorn, co-authors of one of the previous versions of the PFMT review (Hay-Smith 2002).

# REFERENCES

## References to studies included in this review

**Aksac 2003** *{published data only}*

Aksac B, Aki S, Karan A, Yalcin O, Isikoglu M. Biofeedback and pelvic floor exercises for the rehabilitation of urinary stress incontinence. *Gynecologic and Obstetric Investigation* 2003;**56**(1):23-7. [sr-incont16421]

**Asklund 2017** *{published data only}*

* Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a randomized controlled trial. *Neurourology and Urodynamics* 2017;**36**(5):1369-76. [DOI: 10.1002/nau.23116; NCT01848938; sr-incont73594]

Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Treatment of stress urinary incontinence via a smartphone application: a randomised controlled trial. *Neurourology and Urodynamics* 2015;**34**(S3):S40-2. [SR-INCONT68764]

Hoffman V, Soderstrom L, Samuelsson E. Self-management of stress urinary incontinence via a mobile app: two-year follow-up of a randomized controlled trial. *Acta Obstetricia et Gynecologica Scandinavica* 2017;**96**(10):1180-7. [SR-INCONT76865]

Lindh A, Sjostrom M, Stenlund H, Samuelsson E. Non-face-to-face treatment of stress urinary incontinence: predictors of success after 1 year. *International Urogynecology Journal* 2016;**27**(12):1857-65. [SR-INCONT76609
]

Sjostrom M, Lindholm L, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a cost-effectiveness analysis. *Journal of Medical Internet Research* 2017;**19**(5):e154. [SR-INCONT76798]

**Bertotto 2017** *{published data only}*

Bertotto A, Schvartzman R, Wender MCO. Effect of electromyographic biofeedback as an add-on to pelvic floor muscle exercises on neuromuscular outcomes and quality of life in postmenopausal women with stress urinary incontinence: a randomized controlled trial. *Neurourology and Urodynamics* 2017;**36**(8):2142-7. [NCT02275728; SR-INCONT76926]

**Beuttenmuller 2010** *{published data only}*

Beuttenmuller L, Cader SA, Macena RHM, Araujo NDS, Nunes EFC, Dantas EHM. Muscle contraction of the pelvic floor and quality of life of women with stress urinary incontinence who underwent kinesitherapy. *Fizjoterapia* 2010;**18**(1):35-41. [SR-INCONT43326]

**Bidmead 2002** *{published data only}*

* Bidmead J, Mantle J, Cardozo L, Hextall A, Boos K. Home electrical stimulation in addition to conventional pelvic floor exercises: a useful adjunct or expensive distraction? (abstract number 68). *Neurourology and Urodynamics* 2002;**21**(4):372-3. [sr-incont14546]

Parsons M, Mantle J, Cardozo L, Hextall A, Boos K, Bidmead J. A single blind, randomised, controlled trial of pelvic floor muscle training with home electrical stimulation in the treatment of urodynamic stress incontinence (abstract number 296). 34th Annual Meeting of the International Continence Society; 2004 Aug 25-27; Paris, France. 2004.

**Burgio 1998** *{published data only}*

Burgio KL, Locher JL, Goode PS. Combined behavioral and drug therapy for urge incontinence in older women. *Journal of the American Geriatrics Society* 2000;**48**(4):370-4.

* Burgio KL, Locher JL, Goode PS, Hardin JM, McDowell BJ, Dombrowski M, et al. Behavioral vs drug treatment for urge urinary incontinence in older women. *JAMA* 1998;**280**(23):1995-2000.

Burgio KL, Locher JL, Roth DL, Goode PS. Psychological improvements associated with behavioral and drug treatment of urge incontinence in older women. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 2001;**56**(1):46-51.

Goode PS, Burgio KL, Locher JL, Umlauf MG, Lloyd D, Roth DL. Urodynamic changes associated with behavioral and drug treatment of urge incontinence in older women. *Journal of the American Geriatrics Society* 2002;**50**(5):808-16.

**Burns 1993** *{published data only}*

Burns PA, Nochajksi TH, Pranikoff K. Factors discriminating between genuine stress and mixed incontinence. *Journal of the American Academy of Nurse Practitioners* 1992;**4**(1):15-21.

Burns PA, Pranikoff K, Nochajski T, Desotelle P, Harwood MK. Treatment of stress incontinence with pelvic floor exercises and biofeedback. *Journal of the American Geriatrics Society* 1990;**38**(3):341-4.

* Burns PA, Pranikoff K, Nochajski TH, Hadley EC, Levy KJ, Ory MG. A comparison of effectiveness of biofeedback and pelvic muscle exercise treatment of stress incontinence in older community dwelling women. *Journal of Gerontology* 1993;**48**(4):M167-74.

**Bø 1999** *{published data only}*

Bø K, Talseth T, Home I. Single blind, randomised controlled trial of pelvic floor exercises, electrical stimulation, vaginal cones, and no treatment in management of genuine stress incontinence in women. *BMJ* 1999;**318**(7182):487-93. [sr-incont6654]

**Carneiro 2010** *{published data only}*

Carneiro EF, Araujo Ndos S, Beuttenmull L, Vieira PC, Cader SA, Cader SA, et al. The anatomical-functional characteristics of the pelvic floor and quality of life of women with stress urinary incontinence subjected to perineal exercises [Las caracteristicas anatomofuncionales del suelo pelvico y la calidad de vida de mujeres con incontinencia urinaria de esfuerzo sometidas a ejercicios perineales]. *Actas Urologicas Espanolas* 2010;**34**(9):788-93. [SR-INCONT40377]

**Castro 2008** *{published data only}*

Castro RA, Arruda RM, Zanetti MR, Santos PD, Sartori MG, Girao MJ. Single-blind, randomized, controlled trial of pelvic floor muscle training, electrical stimulation, vaginal cones, and no active treatment in the management of stress urinary incontinence. *Clinics (Sao Paulo, Brazil)* 2008;**63**(4):465-72. [sr-incont27760]

**Celiker Tosun 2015** *{published data only}*

Celiker Tosun O, Mutlu EK, Ergenoglu AM, Yeniel AO, Tosun G, Malkoc M, et al. Are symptoms of urinary incontinence reduced in patients when the muscles of the pelvic floor are strengthened through PFM training? What is the relationship between symptoms and strength [abstract 409]. 43rd annual meeting of the International Continence Society; 2013 Aug 27-31; Barcelona, Spain. 2013. [SR-INCONT67955]

Celiker Tosun O, Mutlu EK, Ergenoglu AM, Yeniel AO, Tosun G, Malkoc M, et al. Determining the effect of PFM exercises in UI treatment on PFM strength by non-invasive method, randomised trail [abstract 582]. 43rd annual meeting of the International Continence Society; 2013 Aug 27-31; Barcelona, Spain. 2013. [SR-INCONT50616]

* Celiker Tosun O, Mutlu EK, Ergenoglu AM, Yeniel AO, Tosun G, Malkoc M, et al. Does pelvic floor muscle training abolish symptoms of urinary incontinence? A randomized controlled trial. *Clinical Rehabilitation* 2015;**29**(6):525-37. [DOI: 10.1177/0269215514546768; SR-INCONT68112]

**Diokno 2010** *{published data only}*

Diokno AC, Ocampo MS Jr, Ibrahim IA, Karl CR, Lajiness MJ, Hall SA. Group session teaching of behavioral modification program (BMP) for urinary incontinence: a randomized controlled trial among incontinent women. *International Urology and Nephrology* 2010;**42**(2):375-81. [SR-INCONT39469]

**Ferreira 2014** *{published data only}*

Ferreira S, Ferreira M, Carvalhais A, Santos PC, Rocha P, Brochado G. Reeducation of pelvic floor muscles in volleyball athletes. *Revista da Associacao Medica Brasileira* 2014;**60**(5):428-33. [10.1590/1806-9282.60.05.010; SR-INCONT66745]

**Firra 2013** *{published data only}*

* Firra J, Thompson M, Smith SS. Paradoxical findings in the treatment of predominant stress and urge incontinence: a pilot study with exercise and electrical stimulation. *Journal of Women's Health Physical Therapy* 2013;**37**(3):113-23. [DOI: 10.1097/JWH.0000000000000003; SR-INCONT61836]

Firra JC. Effects of Treatment of Urinary Incontinence in Women: Exercise or Electrical Stimulation or Both [PhD thesis]. Denton (TX): Texas Woman's University, 2008:1-206. [sr-incont47108]

**Henalla 1989** *{published data only}*

Henalla SM, Hutchins CJ, Robinson P, MacVicar J. Non-operative methods in the treatment of female genuine stress incontinence of urine. *Journal of Obstetrics and Gynaecology* 1989;**9**(3):222-5. [sr-incont2637]

**Henalla 1990** *{published data only}*

Henalla SM, Millar DR, Wallace KJ. Surgical versus conservative management for post-menopausal genuine stress incontinence of urine (abstract number 87). *Neurourology and Urodynamics* 1990;**9**(4):436-7. [sr-incont5117]

**Hofbauer 1990** *{published data only}*

* Hofbauer VJ, Preisinger F, Nurnberger N. The value of physical therapy in genuine female stress incontinence [Der Stellenwert der Physikotherapie bei der weiblichen genuinen Stress-inkontinenz]. *Zeitschrift fur Urologie und Nephrologie* 1990;**83**(5):249-54. [SR-INCONT350]

Preisinger E, Hofbauer J, Nurnberger N, Sadil S, Schneider B. Possibilities of physiotherapy for urinary stress incontinence [Möglichkeiten der physikalischen therapie bei Streßinkontinenz]. *Zeitschrift fur physikalische Medizin, Balneologie, med. Klimatologie* 1990;**19**(2):75-9. [0720-9762; SR-INCONT6314]

**Kargar Jahromi 2013** *{published data only}*

Kargar Jahromi M, Talebizadeh M, Mirzaei M. The effect of pelvic muscle exercises on urinary incontinency and self-esteem of elderly females with stress urinary incontinency, 2013. *Global Journal of Health Science* 2015;**7**(2):71-9. [DOI: 10.5539/gjhs.v7n2p71; SRINCONT66372]

**Kim 2007** *{published data only}*

Kim H, Suzuki T, Yoshida Y, Yoshida H. Effectiveness of multidimensional exercises for the treatment of stress urinary incontinence in elderly community-dwelling Japanese women: a randomized, controlled, crossover trial. *Journal of the American Geriatrics Society* 2007;**55**(12):1932-9. [sr-incont26012]

**Kim 2011a** *{published data only}*

Kim H, Yoshida H, Suzuki T. Effects of exercise treatment with or without heat and steam generating sheet on urine loss in community-dwelling Japanese elderly women with urinary incontinence. *Geriatrics & Gerontology International* 2011;**11**(4):452-9. [SR-INCONT42661]

**Kim 2011b** *{published data only}*

Kim H, Yoshida H, Suzuki T. Exercises treatment to reduce the urine leakage in elderly community-dwelling Japanese women with stress, urge, and mixed urinary incontinence [abstract 772]. 39th Annual Meeting of the International Continence Society (ICS); 2009 Sept 29 - Oct 3, San Francisco, CA. 2009. [SR-INCONT35628]

* Kim H, Yoshida H, Suzuki T. The effects of multidimensional exercise treatment on community-dwelling elderly Japanese women with stress, urge, and mixed urinary incontinence: a randomized controlled trial. *International Journal of Nursing Studies* 2011;**48**(10):1165-72. [SR-INCONT42657]

**Lagro-Janssen 1991a** *{published data only}*

Lagro-Janssen AL, Debruyne FM, Smits AJ, van Weel C. The effects of treatment of urinary incontinence in general practice. *Family Practice* 1992;**9**(3):284-9.

Lagro-Janssen T, van Weel C. Long-term effect of treatment of female incontinence in general practice. *British Journal of General Practice* 1998;**48**(436):1735-8.

* Lagro-Janssen TL, Debruyne FM, Smits AJ, van Weel C. Controlled trial of pelvic floor exercises in the treatment of urinary stress incontinence in general practice. *British Journal of General Practice* 1991;**41**(352):445-9. [sr-incont232]

**Leong 2015** *{published data only}*

Leong BS, Mok NW. Effectiveness of a new standardised Urinary Continence Physiotherapy Programme for community-dwelling older women in Hong Kong. *Hong Kong Medical Journal* 2015;**21**(1):30-7. [DOI: 10.12809/hkmj134185; SR-INCONT66405]

**McLean 2013** *{published data only}*

McLean L, Varette K, Gentilcore-Saulnier E, Harvey MA, Baker K, Sauerbrei E. Pelvic floor muscle training in women with stress urinary incontinence causes hypertrophy of the urethral sphincters and reduces bladder neck mobility during coughing. *Neurourology and Urodynamics* 2013;**32**(8):1096-102. [DOI: 10.1002/nau.22343; SR-INCONT49427]

**Miller 1998** *{published data only}*

Miller JM, Ashton-Miller JA, DeLancey JO. A pelvic muscle precontraction can reduce cough-related urine loss in selected women with mild SUI. *Journal of the American Geriatrics Society* 1998;**46**(7):870-4. [sr-incont6659]

**Pereira 2011** *{published data only}*

Nascimento-Correia G, Santos-Pereira V, Tahara N, Driusso P. Effects of pelvic floor muscle training on quality of life of a group of women with urinary incontinence: randomized controlled trial [Efectos del fortalecimiento del suelo pelvico en la calidad de vida de un grupo de mujeres con incontinencia urinaria: estudio aleatorizado controlado]. *Actas Urologicas Espanolas* 2012;**36**(4):216-21. [sr-incont44721]

* Pereira VS, Correia GN, Driusso P. Individual and group pelvic floor muscle training versus no treatment in female stress urinary incontinence: a randomized controlled pilot study. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2011;**159**(2):465-71. [SR-INCONT42734]

Pereira VS, de Melo MV, Correia GN, Driusso P. Long-term effects of pelvic floor muscle training with vaginal cone in post-menopausal women with urinary incontinence: a randomized controlled trial. *Neurourology and Urodynamics* 2013;**32**(1):48-52. [SR-INCONT46703]

Pereira VS, de Melo MV, Correia GN, Driusso P. Vaginal cone for postmenopausal women with stress urinary incontinence: randomized, controlled trial. *Climacteric* 2012;**15**(1):45-51. [SR-INCONT42979]

**Sar 2009** *{published data only}*

Sar D, Khorshid L. The effects of pelvic floor muscle training on stress and mixed urinary incontinence and quality of life. *Journal of Wound, Ostomy, and Continence Nursing* 2009;**36**(4):429-35. [SR-INCONT32103]

**Solberg 2016** *{published data only}*

Solberg M, Alraek T, Mdala I, Klovning A. A pilot study on the use of acupuncture or pelvic floor muscle training for mixed urinary incontinence. *Acupuncture in Medicine* 2016;**34**(1):7-13. [DOI: 10.1136/acupmed-2015-010828; SR-INCONT70460]

**Sran 2016** *{published data only}*

Dumoulin C, Sran M, Lieblich P, Wilson P. Physiotherapy significantly reduces leakage in postmenopausal women with osteoporosis and urinary incontinence: result of a parallel RCT [abstract 130]. *Neurourology and Urodynamics* 2011;**30**(6):985.

* Sran M, Mercier J, Wilson P, Lieblich P, Dumoulin C. Physical therapy for urinary incontinence in postmenopausal women with osteoporosis or low bone density: a randomized controlled trial. *Menopause* 2016;**23**(3):286-93. [DOI: 10.1097/GME.0000000000000594; SR-INCONT70407]

**Wells 1999** *{published and unpublished data}*

Wells TJ, Mayor RD, Brink CA, Brown R. Pelvic floor muscle exercise: a controlled clinical trial (unpublished manuscript as supplied 3 November 2008). Manuscript on file (unpublished work; however, portions were presented at the 16th annual scientific meeting of the American Urogynecologic Society; 1999 Oct 12-15; Seattle (WA)). [sr-incont26977]

**Yoon 2003** *{published data only}*

Yoon HS, Song HH, Ro YJ. A comparison of effectiveness of bladder training and pelvic muscle exercise on female urinary incontinence. *International Journal of Nursing Studies* 2003;**40**(1):45-50. [sr-incont15700]

## References to studies excluded from this review

**Abdulaziz 2012** *{published data only}*

Abdulaziz K, Hasan T. Role of pelvic floor muscle therapy in obese perimenopausal females with stress incontinence: a randomized control trial. Internet Journal of Gynecology and Obstetrics 2012; Vol. 16, issue 2. [1528-8439]

**Albers-Heitner 2008** *{published data only}*

Albers-Heitner P, Berghmans B, Joore M, Lagro-Janssen T, Severens J, Nieman F, et al. The effects of involving a nurse practitioner in primary care for adult patients with urinary incontinence: the PromoCon study (Promoting Continence). *BMC Health Services Research* 2008;**8**:84. [DOI: 10.1186/1472-6963-8-84]

**Albers-Heitner 2013** *{published data only}*

Albers-Heitner P, Winkens R, Berghmans B, Joore M, Nieman F, Severens J, et al. Consumer satisfaction among patients and their general practitioners about involving nurse specialists in primary care for patients with urinary incontinence. *Scandinavian Journal of Caring Sciences* 2013;**27**(2):253-9. [DOI: 10.1111/j1471-6712.2012.01023.x; ISRCTN62722772; SR-INCONT48048]

**Alves 2014** *{published data only}*

Alves F, Botelho S, Vasconcelos D, Marques J, Pereira L, Palma P, et al. A group pelvic floor muscle training programme increases muscular contractility in postmenopausal women:

a randomized controlled trial. 43rd Annual Meeting of the International Continence Society; 2013 Aug 27-31; Barcelona, Spain. 2013. [SR-INCONT49729]

* Alves FK, Riccetto C, Adami DB, Marques J, Pereira LC, Palma P, et al. A pelvic floor muscle training program in postmenopausal women: a randomized controlled trial. *Maturitas* 2014;**81**(2):300-5. [DOI: 10.1016/j.maturitas.2015.03.006; SR-INCONT66852]

**Bernier 2008a** *{published data only}*

Bernier F. Pelvic Floor Muscle Retraining: Quantitative, Experimental, Randomized Pilot Study [PhD thesis]. Charlottesville, VA: University of Virginia, 2008.

**Bernier 2008b** *{published data only}*

Bernier F. Pelvic floor muscle retraining: quantitative, experimental, randomized pilot study. 39th Annual Conference of the Society of Urologic Nurses and Associates (SUNA), 2008 Oct 3-6; Philadelphia (PA) 2008.

**Beuttenmuller 2011** *{published data only}*

Beuttenmuller L, Cader SA, Macena RHM, Araujo NDS, Nunes EFC, Dantas EHM. Floor muscles contraction in women with stress urinary incontinence underwent to exercises and electric stimulation therapy: a randomized study. *Fisioterapia e Pesquisa* 2011;**18**(3):210-6. [SR-INCONT47117]

**Burgio 2002** *{published data only}*

Burgio KL, Goode PS, Locher JL, Umlauf MG, Roth DL, Richter HE, et al. Behavioral training with and without biofeedback in the treatment of urge incontinence in older women. A randomized controlled trial. *JAMA* 2002;**288**(18):2293-9.

**Chang 2011** *{published data only}*

Chang KK, Wong TK, Wong TH, Leung AW, Chung JW. Effect of acupressure in treating urodynamic stress incontinence: a randomized controlled trial. *American Journal of Chinese Medicine* 2011;**39**(6):1139-59.

**Dugan 2013** *{published data only}*

Dugan SA, Lavender MD, Hebert-Beirne J, Brubaker L. A pelvic floor fitness program for older women with urinary symptoms: a feasibility study. *PM & R : the Journal of Injury, Function, and Rehabilitation* 2013;**5**(8):672-6. [DOI: 10.1016/j.pmrj.2013.03.009; SR-INCONT48556]

**Felicissimo 2010** *{published data only}*

Felicissimo MF, Carneiro MM, Saleme CS, Pinto RZ, da Fonseca AM, da Silva-Filho AL. Intensive supervised versus unsupervised pelvic floor muscle training for the treatment of stress urinary incontinence: a randomized comparative trial. *International Urogynecology Journal* 2010;**21**(7):835-40.

**Ferreira 2011a** *{published data only}*

Ferreira M, Santos P, Duarte J. Influence of home-based and supervised exercise programs on women stress urinary incontinence: a randomized trial. *Physiotherapy* 2011;**97**(Suppl S1):eS339-40.

**Ferreira 2011b** *{published data only}*

Ferreira M, Pitanguy M, Vasconcelos ECLM, Antonio FI. Evaluation of urinary incontinence reports and pelvic floor muscle knowledge among women participating in a pelvic floor educational group. *Physiotherapy* 2011;**97**(Suppl S1):eS337-8.

**Ghaderi 2016** *{published data only}*

Ghaderi F, Mohammadi K, Sasan RA, Kheslat SN, Oskouel AE. Effects of stabilization exercises focusing on pelvic floor muscles on low back pain and urinary incontinence in women. *Urology* 2016;**93**:50-4. [IRCT2014081218760N1; SR-INCONT72138]

**Ghoniem 2005** *{published data only}*

Ghoniem GM, Van Leeuwen JS, Elser DM, Freeman RM, Zhao YD, Yalcin I, et al. A randomized controlled trial of duloxetine alone, pelvic floor muscle training alone, combined treatment and no active treatment in women with stress urinary incontinence. *Journal of Urology* 2005;**173**(5):1647-53.

**Ghroubi 2015** *{published data only}*

Ghroubi S, Elleuch W, Mahersi S, Feki H, Elleuch MH. Effect of adding pelvic floor muscle training to weight loss intervention on urinary incontinence in overweight women: a randomized controlled trial (abstract number CO40-005-e). *Annals of Physical and Rehabilitation Medicine* 2015;**58S**(Suppl 1):e132. [SR-INCONT71127]

**Goode 2003** *{published data only}*

Goode PS, Burgio KL, Locher JL, Roth DL, Umlauf MG, Richter HE, et al. Effect of behavioral training with or without pelvic floor electrical stimulation on stress incontinence in women. A randomized controlled trial. *JAMA* 2003;**290**(3):345-52.

**Kang 2015** *{published data only}*

Kang H, Hong G-RS. Effect of muscle strength training on urinary incontinence and physical function: a randomized controlled trial in long-term care facilities. *Journal of Korean Academy of Nursing* 2015;**45**(1):35-45. [SR-INCONT66713]

**Kumari 2008** *{published data only}*

Kumari S, Jain V, Mandal AK, Singh A. Behavioral therapy for urinary incontinence in India. *International Journal of Gynaecology and Obstetrics* 2008;**103**(2):125-30.

**Lapointe 2014** *{published data only}*

Lapointe A, Tannenbaum C. Should octogenarians (80 years and over) be included in health education randomized controlled trials?. *Journal of the American Geriatrics Society* 2014;**62**(Suppl s1):S37. [DOI: 10.1111/jgs.12870; SR-INCONT61851]

**McLean 2015** *{published data only}*

McLean L, Varette K, Baker K, Harvey M, Brison R, Day A. Women with less severe stress urinary incontinence symptoms are most likely to show a complete cure on a standardized pad test after performing a 12 week exercise program. *Neurourology and Urodynamics* 2015;**34**(S3):S42-3. [DOI: 10.1002/nau.22830; SR-INCONT71184]

**Miller 2009** *{published data only}*

Miller J, Park L, Tolbert M, Ashton-Miller J, DeLancey J. Short term response to Knack therapy (no dedicated muscle strengthening) for treatment of incontinence (abstract number 548). 39th Annual Meeting of the International Continence Society, 2009 Sept 29 - Oct 3, San Francisco, CA. 2009.

**NCT02001714** *{published data only}*

* NCT02001714, Diokno AC. Group learning achieves decreased incidents of lower urinary symptoms (Gladiolus). clinicaltrials.gov/show/NCT02001714 Date first received: 5 December 2013. [NCT02001714; SR-INCONT60912]

Newman D, Diokno A, Low L, Goode P, Burgio K, Griebling T, et al. A multi-site prospective randomized controlled trial of group administered behavioral treatment in reducing urinary incontinence in older adult women. *Neurourology and Urodynamics* 2017;**36**(Suppl S3):S390-1. [SR-INCONT77531]

**NTR1114** *{published data only}*

NTR1114, Hazewinkel MH. Pelvic physiotherapy after surgical treatment for cervical cancer [A randomized controlled trial to evaluate the effect on pelvic floor function of pelvic physiotherapy following radical hysterectomy in early stage cervical cancer]. trialregister.nl/trialreg/admin/rctview.asp?TC=1114 Date first received: 5 November 2007. [NTR1114; sr-incont34192]

**Ramsay 1990** *{published data only}*

Ramsay IN, Thou M. A randomised, double blind, placebo controlled trial of pelvic floor exercises in the treatment of genuine stress incontinence (abstract number 59). *Neurourology and Urodynamics* 1990;**9**(4):398-9.

**RBR-3fgwc7** *{published data only}*

RBR-3fgwc7, Knorst K. Treatment through electrical stimulation and exercises for urinary incontinence. Pelvic floor training in urinary incontinence comparing techniques treatment by intravaginal electrical stimulation and kinesiotherapy. ensaiosclinicos.gov.br/rg/RBR-3fgwc7/ Date first received: 12 January 2016. [RBR-3fgwc7; SR-INCONT76220]

**Rutledge 2012** *{published data only}*

Rutledge T, Lee S, Rogers R, Muller C. A pilot randomized controlled trial to evaluate pelvic floor muscle training for urinary incontinence among gynecologic cancer survivors (abstract number 282). *Gynecologic Oncology* 2012;**125**(Suppl 1):S117.

**Sjostrom 2015** *{published data only}*

Bjork AB, Sjostrom M, Johansson EE, Samuelsson E, Umefjord G. Women's experiences of internet-based or postal treatment for stress urinary incontinence. *Qualitative Health Research* 2014;**24**(4):489-93. [SR-INCONT60695]

Lindh A, Sjöström M, Stenlund H, Samuelsson E. Non-face-to-face treatment of stress urinary incontinence: predictors of success after 1 year (abstract number 521). *Neurourology and Urodynamics* 2015;**34**(S3):S443-4. [SR-INCONT68733]

Sjostrom M, Umefjord G, Lindholm L, Samuelson E. Cost-effectiveness of an internet-based treatment program for stress urinary incontinence. *Neurourology and Urodynamics* 2015;**34**(3):244-50. [SRINCONT66710]

Sjostrom M, Umefjord G, Stenlund H, Carlbring P, Andersson G, Samuelsson E. Internet-based treatment of stress urinary incontinence: a randomised controlled study with focus on pelvic floor muscle training. *BJU International* 2013;**112**(3):362-72. [SR-INCONT48423]

* Sjöström M, Umefjord G, Stenlund H, Carlbring P, Andersson G, Samuelsson E. Internet-based treatment of stress urinary incontinence: 1- and 2-year results of a randomized controlled trial with a focus on pelvic floor muscle training. *BJU International* 2015;**116**(6):955-64. [SR-INCONT70287]

**Tajiri 2014** *{published data only}*

Tajiri K, Huo M, Maruyama H. Effects of co-contraction of both transverse abdominal muscle and pelvic floor muscle exercises for stress urinary incontinence: a randomized controlled trial. *Journal of Physical Therapy Science* 2014;**26**(8):1161-3. [SR-INCONT66689]

**Talley 2017** *{published data only}*

Talley KM, Wyman JF, Bronas U, Olson-Kellogg BJ, McCarthy TC. Defeating urinary incontinence with exercise training: results of a pilot study in frail older women. *Journal of the American Geriatrics Society* 2017;**65**(6):1321-7. [SR-INCONT74510]

**van Leeuwen 2004** *{published data only}*

van Leeuwen JHS, Elser D, Freeman R, Ghoniem G, Zhao Y, Yalcin I, et al. Controlled trial of duloxetine alone, pelvic floor muscle training alone, combined treatment in women with stress urinary incontinence (SUI) (abstract number 197). *European Urology Supplements* 2004;**3**(2):52.

**Yang 2012** *{published data only}*

Yang EJ, Lim JY, Rah UW, Kim YB. Effect of a pelvic floor muscle training program on gynecologic cancer survivors with pelvic floor dysfunction: a randomized controlled trial. *Gynecologic Oncology* 2012;**125**(3):705-11.

**Yoon 1999** *{published data only}*

Yoon HN, Hong JY, Choi YH, Back SH. The effect of pelvic floor muscle exercises on genuine stress incontinence among Korean women: focusing on its effects on the quality of life. 29th Annual Meeting of the International Continence Society, 1999 Aug 23-26; Denver (CO). 1999.

## References to studies awaiting assessment

**Bali 2016** *{published data only}*

Bali P, Mahalingam G, Bala K. Effectiveness of Kegal exercise on women with urinary incontinence. *International Journal of Nursing Education* 2016;**8**(3):155-62. [SR-INCONT73391]

**Zhang 2015** *{published data only}*

Zhang Y, Yi NH, Wu L, Fang JX, Tang N. Effect of pelvic floor rehabilitation on stress urinary incontinence of elderly female (abstract number P19). *Journal of the American Geriatrics Society* 2015;**63**(Suppl S2):S328. [SR-INCONT72732]

## References to ongoing studies

**NCT03097549** *{published data only}*

NCT03097549, Samuelsson E. Mobile app-treatment of mixed and urgency urinary incontinence in women. clinicaltrials.gov/show/NCT03097549 Date first received: 31 March 2017. [NCT03097549; SR-INCONT76074]

## Additional references

**Avery 2004**

Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. *Neurourology and Urodynamics* 2004;**23**(4):322-30.

**Ayeleke 2015**

Ayeleke RO, Hay-Smith EJC, Omar MI. Pelvic floor muscle training added to another active treatment versus the same active treatment alone for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2015, Issue 11. [DOI: 10.1002/14651858.CD010551.pub3]

**Balmforth 2004**

Balmforth J, Bidmead J, Cardozo L, Hextall A, Kelvin B, Mantle J. Raising the tone: A prospective observational study evaluating the effect of pelvic floor muscle training on bladder neck mobility and associated improvement in stress urinary incontinence (abstract number 111). *Neurourology and Urodynamics* 2004;**23**(5-6):553-4.

**Berghmans 1998**

Berghmans LC, Hendriks HJ, Bø K, Hay-Smith EJ, de Bie RA, van Waalwijk van Doorn ES. Conservative treatment of stress urinary incontinence in women: a systematic review of randomized clinical trials. *British Journal of Urology* 1998;**82**(2):181-91.

**Berghmans 2000**

Berghmans LC, Hendriks HJ, de Bie RA, van Waalwijk van Doorn ES, Bø K, van Kerrebroeck PhE. Conservative treatment of urge urinary incontinence in women: a systematic review of randomized clinical trials. *BJU International* 2000;**85**(3):254-63.

**Boutron 2008**

Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine* 2008;**148**(4):295-309.

**Brookes 2004**

Brookes ST, Donovan JL, Wright M, Jackson S, Abrams P. A scored form of the Bristol Female Lower Urinary Tract Symptoms questionnaire: data from a randomized controlled trial of surgery for women with stress incontinence. *American Journal of Obstetrics and Gynecology* 2004;**191**(1):72-82.

**Broome 2003**

Broome BA. The impact of urinary incontinence on self-efficacy and quality of life. *Health and Quality of Life Outcomes* 2003;**1**(1):35.

**Bushnell 2005**

Bushnell DM, Martin ML, Summers KH, Svihra J, Lionis C, Patrick DL. Quality of life of women with urinary incontinence: cross-cultural performance of 15 language versions of the I-QOL. *Quality of Life Research* 2005;**14**(8):1901-13.

**Bø 1994**

Bø K. Reproducibility of instruments designed to measure subjective evaluation of female stress urinary incontinence. *Scandinavian Journal of Urology and Nephrology* 1994;**28**(1):97-100.

**Bø 1995**

Bø K. Pelvic floor muscle exercise for the treatment of stress urinary incontinence: an exercise physiology perspective. *International Urogynecology Journal and Pelvic Floor Dysfunction* 1995;**6**(5):282-91.

**Bø 1996**

Bø K. Physiotherapy to treat genuine stress incontinence. *International Continence Survey* 1996;**6**:2-8.

**Bø 2001**

Bø K, Lilleas F, Talseth T, Hedland H. Dynamic MRI of the pelvic floor muscles in an upright sitting position. *Neurourology and Urodynamics* 2001;**20**(2):167-74.

**Bø 2004**

Bø K. Pelvic floor muscle training is effective in treatment of female stress urinary incontinence, but how does it work?. *International Urogynecology Journal and Pelvic Floor Dysfunction* 2004;**15**(2):76-84.

**Bø 2007**

Bø K, Bergmans B, Morkved S, Van Kampen M. Evidence-Based Physical Therapy for the Pelvic Floor. Edinburgh: Harrison H, 2007.

**Bø 2017**

Bø K, Frawley H, Haylen BT, Abramov Y, Almeida FG, Berghmans B, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for the conservative and nonpharmacological management of female pelvic floor dysfunction. *International Urogynecology Journal* 2017;**28**(2):191-213.

**Carrière 2006**

Carrière B. The Pelvic Floor. Stuttgart: Georg Thieme Verlag, 2006.

**Chang 1984**

Chang S. Internal Exercise. Oslo: Stuyvesant Publishing Company, 1984.

**Clement 2013**

Clement KD, Lapitan MCM, Omar MI, Glazener CMA. Urodynamic studies for management of urinary incontinence in children and adults. *Cochrane Database of Systematic Reviews* 2013, Issue 10. [DOI: 10.1002/14651858.CD003195.pub3]

**Constantinou 1982**

Constantinou CE, DE Govan. Spatial distribution and timing of transmitted and reflexly generated urethral pressures in healthy women. *Journal of Urology* 1982;**127**(5):964-9. [PUBMED: 7201031]

**Coyne 2012**

Coyne KS, Kvasz M, Ireland AM, Milsom I, Kopp ZS, Chapple CR. Urinary incontinence and its relationship to mental health and health-related quality of life in men and women in Sweden, the United Kingdom, and the United States. *European Urology* 2012;**61**(1):88-95.

**de Groat 2001**

de Groat WC, Yoshimura N. Pharmacology of the lower urinary tract. *Annual Review of Pharmacology and Toxicology* 2001;**41**:691-721.

**de Kruif 1996**

de Kruif YP, van Wegen EE. Pelvic floor muscle exercise therapy with myofeedback for women with stress urinary incontinence: a meta-analysis. *Physiotherapy* 1996;**82**(2):107-13.

**DeLancey 1988**

DeLancey JO. Structural aspects of urethrovesical function in the female. *Neurourology and Urodynamics* 1988;**7**(6):509-19.

**Derogatis 1974**

Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins symptom check-list (HSCL): a self-report symptom inventory. *Behavioral Science* 1974;**19**(1):1-15.

**Derogatis 1983**

Derogatis LR. The SCL-90-R Administration, Scoring, and Procedures Manual II. Towson, MD: Clinical Psychometric Research, 1983.

**Donovan 2005**

Donovan J, Bosch R, Gotoh M, Jackson S, Naughton M, Radley S, et al. Symptom and quality of life assessment. Incontinence. Volume 1 Basics & Evaluation. Edition 2005. Plymouth: Health Publication Ltd, 2005:519-84.

**Dumoulin 2007**

Dumoulin C, Peng Q, Stodkilde-Jorgensen H, Shishido K, Constantinou C. Changes in levator ani anatomical configuration following physiotherapy in women with stress urinary incontinence. *Journal of Urology* 2007;**178**(3 Pt 1):970-7; quiz 1129.

**Dumoulin 2011**

Dumoulin C, Glazener C, Jenkinson D. Determining the optimal pelvic floor muscle training regimen for women with stress urinary incontinence. *Neurourology and Urodynamics* 2011;**30**(5):746-53.

**Dumoulin 2017**

Dumoulin C, Adewuyi T, Booth J, Bradley C, Burgio K, Hagen S, et al. Adult conservative management. In: Abrams P, Cardozo L, Wagg A, Wein A editor(s). Incontinence: 6th International Consultation on Incontinence, Tokyo, September 2016. Bristol, UK: International Continence Society (ICS) and International Consultation on Urological Diseases (ICUD), 2017:1443-628.

**Fantl 1996**

Fantl J, Newman DK, Colling J. Urinary incontinence in adults: acute and chronic management: 1996 update. AHCPR clinical practice guidelines No. 2. Rockville (MD): Agency for Health Care Policy and Research (AHCPR), Public Health Service, US Department of Health and Human Services; 1996 Mar. AHCPR Report No.: 96-0682.

**Fedorkow 1993**

Fedorkow DM. Nonsurgical management of stress urinary incontinence. *Journal SOGC : Journal of the Society of Obstetricians and Gynaecologists of Canada* 1993;**15**(6):695-705.

**Ferguson 2002**

Ferguson D, Aaron SD, Guyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002;**325**:652-4.

**Godec 1975**

Godec C, Cass AS, Ayala GF. Bladder inhibition with functional electrical stimulation. *Urology* 1975;**6**(6):663-6.

**Griffin 1994**

Griffin C, MC Dougherty, Yarandi H. Pelvic muscles during rest: responses to pelvic muscle exercise. *Nursing Research* 1994;**43**(3):164-7.

**Guyatt 2013a**

Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines: 12. Preparing summary of findings tables - binary outcomes. *Journal of Clinical Epidemiology* 2013;**66**(2):158-72.

**Guyatt 2013b**

Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles - continuous outcomes. *Journal of Clinical Epidemiology* 2013;**66**(2):173-83.

**Hannestad 2000**

Hannestad YS, Rortveit G, Sandvik H, Hunaskaar S. A community-based epidemiological survey of female urinary incontinence: the Norwegian EPINCONT study. Epidemiology of Incontinence in the County of Nord-Trøndelag. *Journal of Clinical Epidemiology* 2000;**53**(11):1150-7.

**Hay-Smith 2002**

Hay-Smith EJ, Bø K, Berghmans LC, Hendriks HJ, de Bie RA, van Waalwijk van Doorn ES. Pelvic floor muscle training for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2002, Issue 1. [DOI: 10.1002/14651858.CD001407]

**Hay-Smith 2011**

Hay-Smith EJ, Herderschee R, Dumoulin C, Herbison GP. Comparisons of approaches to pelvic floor muscle training for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2011, Issue 12. [DOI: 10.1002/14651858.CD009508; CD009508]

**Haylen 2010**

Haylen BT, de Ridder D, Freeman RM, Swift SE, Berghmans B, Lee J, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. *Neurourology and Urodynamics* 2010;**29**(1):4-20.

**Herbert 2005**

Herbert RD, Bø K. Analysis of quality of interventions in systematic reviews. *BMJ* 2005;**331**(7515):507-9.

**Herderschee 2011**

Herderschee R, Hay-Smith EJ, Herbison GP, Roovers JP, Heineman MJ. Feedback or biofeedback to augment pelvic floor muscle training for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2011, Issue 7. [DOI: 10.1002/14651858.CD009252]

**Higgins 2011**

Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

**Hunskaar 1991**

Hunskaar S, Vinsnes A. The quality of life in women with urinary incontinence as measured by the sickness impact profile. *Journal of the American Geriatrics Society* 1991;**39**(4):378-82.

**Hunskaar 2002**

Hunskaar S, Burgio K, Diokno AC, Herzog AR, Hjalmas K, Lapitan MC. Epidemiology and natural history of urinary incontinence (UI). In: Abrams P, Cardozo L, Khoury S editor(s). Incontinence. 2nd Edition. Plymouth: Health Publication Ltd, 2002.

**Imamura 2010**

Imamura M, Abrams P, Bain C, Buckley B, Cardozo L, Cody J, et al. Systematic review and economic modelling of the effectiveness and cost-effectiveness of non-surgical treatments for women with stress urinary incontinence. *Health Technology Assessment (Winchester, England)* 2010;**14**(40):1-506. [DOI: 10.3310/hta14400]

**Jackson 1996**

Jackson S, Donovan J, Brookes S, Eckford S, Swithinbank L, Abrams P. The Bristol Female Lower Urinary Tract Symptoms questionnaire: development and psychometric testing. *British Journal of Urology* 1996;**77**(6):805-12.

**Johannesson 1997**

Johannesson M, O'Conor R, Kobelt-Ngyuen G, Mattiasson M. Willingness to pay for reduced incontinence symptoms. *British Journal of Urology* 1997;**80**:557-62.

**Johnson 1998**

Johnson TM 2nd, Kincade JE, Bernard SL, Busby-Whitehead J, Hertz-Picciotto I, DeFriese GH. The association of urinary incontinence with poor self-rated health. *Journal of the American Geriatrics Society* 1998;**46**(6):693-9.

**Kegel 1948**

Kegel AH. Progressive resistance exercise in the functional restoration of the perineal muscles. *American Journal of Obstetrics and Gynecology* 1948;**56**(2):238-48.

**Kelleher 1997**

Kelleher CJ, Cardozo LD, Khullar V, Salvatore S. A new questionnaire to assess the quality of life of urinary incontinent women. *British Journal of Obstetrics and Gynaecology* 1997;**104**(12):1374-9.

**Kelleher 2013**

Kelleher C, Staskin D, Cherian P, Cotterill N, Coyne K, Kopp Z, et al. Patient-reported outcome assessment. In: Abrams P, Cardozo L, Khoury S, Wein A editor(s). Incontinence. 5th Edition. Paris: Health Publication Ltd, 2013:389-428.

**Lagro-Janssen 1991b**

Lagro-Janssen TL, Debruyne FM, Smits AJ, van Weel C. Controlled trial of pelvic floor exercises in the treatment of urinary stress incontinence in general practice. *British Journal of General Practice* 1991;**41**(352):445-9.

**Lee 1995**

Lee PS, Reid DW, Saltmarche A, Linton L. Measuring the psychosocial impact of urinary incontinence: the York Incontinence Perceptions Scale (YIPS). *Journal of the American Geriatrics Society* 1995;**43**(11):1275-8.

**Lins 2014**

Lins S, Hayder-Beichel D, Kunath F, Cartwright R, Motschall E, Antes G, et al. Pelvic floor muscle training versus other active treatments for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2014, Issue 6. [DOI: 10.1002/14651858.CD007173.pub3]

**Lose 1998**

Lose G, Fantl JA, Victor A, Walter S, Wells TL, Wyman J, et al. Outcome measures for research in adult women with symptoms of lower urinary tract dysfunction. *Neurourology and Urodynamics* 1998;**17**(3):255-62.

**Margolis 2011**

Margolis MK, Vats V, Coyne KS, Kelleher C. Establishing the content validity of the King's Health Questionnaire in men and women with overactive bladder in the US. *Patient* 2011;**4**(3):177-87.

**McClurg 2016**

McClurg D, Pollock A, Campbell P, Hazelton C, Elders A, Hagen S, et al. Conservative interventions for urinary incontinence in women: an Overview of Cochrane systematic reviews. *Cochrane Database of Systematic Reviews* 2016, Issue 9. [DOI: 10.1002/14651858.CD012337]

**Milsom 2013**

Milsom I, Altman D, Cartwright R, Lapitan MC, Nelson R, Sillen U, et al. Epidemiology of urinary incontinence (UI) and other lower urinary tract symptoms (LUTS), pelvic organ prolapse (POP) and anal incontinence (AI). In: Abrams P, Cardozo L, Khoury S, Wein A editor(s). Incontinence: 5th International Consultation on Incontinence. Recommendations of the International Scientific Committee: evaluation and treatment of urinary incontinence, pelvic organ prolapse and faecal incontinence; 2012 Feb 23-25; Paris. Belgium: International Consultation on Urological Diseases (ICUD), 2013:15-107.

**Moher 1998**

Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses?. *Lancet* 1998;**352**(9128):609-13.

**Moher 2001**

Moher M, Schulz KF, Altman DG, for the CONSORT group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;**285**(15):1987-91.

**Morris 1936**

Morris M. Maternity and Post-operative Exercises. London: William Heinemann Ltd, 1936:5-11, 60-65.

**Morrison 1995**

Morrison JF. The excitability of the micturition reflex. *Scandinavian Journal of Urology and Nephrology. Supplementum* 1995;**175**:21-5. [PUBMED: 8771271]

**Papanicolaou 2005**

Papanicolaou S, Hunskaar S, Lose G, Sykes D. Assessment of bothersomeness and impact on quality of life of urinary incontinence in women in France, Germany, Spain and the UK. *BJU International* 2005;**96**(6):831-8.

**Peschers 2001**

Peschers UM, Vodusek DB, Fanger G, Schaer GN, DeLancey JOL, Schuessler B. Pelvic muscle activity in nulliparous volunteers. *Neurourology and Urodynamics* 2001;**20**(3):269-75.

**Pontbriand-Drolet 2012**

Pontbriand-Drolet S, Tang A, Madill S, Bergeron J, Dumoulin C. An MRI study: Do pelvic floor morphological and functional deficit differ in women with stress compared to mixed urinary incontinence? (abstract number 270). *Neurourology and Urodynamics* 2012;**31**(6):1076-8.

**Ramsay 1995**

Ramsay IN, Ali HM, Hunter M, Stark D, Donaldson K. A randomised controlled trial of urodynamic investigations prior to conservative treatment of urinary incontinence in the female. *International Urogynecology Journal* 1995;**6**(5):277-81.

**Reference Manager 2012 [Computer program]**

Thomson Reuters. Reference Manager Professional Edition Version 12. New York: Thomson Reuters, 2012.

**Resnick 1989**

Resnick NM, Yalla SV, Laurino E. The pathophysiology of urinary incontinence among institutionalized elderly persons. *New England Journal of Medicine* 1989;**320**(1):1-7.

**Review Manager 2014 [Computer program]**

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

**Sandvik 2000**

Sandvik H, Seim A, Vanvik A, Hunskaar S. A severity index for epidemiological surveys of female urinary incontinence: comparison with 48-hour pad-weighing tests. *Neurourology and Urodynamics* 2000;**19**(2):137-45.

**Sen 2006**

Sen I, Onaran M, Aksakal N, Acar C, Ozgur Tan M, Acar A, et al. The impact of urinary incontinence on female sexual function. *Advances in Therapy* 2006;**23**(6):999-1008.

**Shemilt 2011**

Shemilt I, Mugford M, Vale L, Craig D and on behalf of the Campbell and Cochrane Economics Methods Group. Searching NHS EED and HEED to inform development of economic commentary for Cochrane intervention reviews [Online]. methods.cochrane.org/economics/workshops (accessed 23 July 2018).

**Sjostrom 2017**

Sjostrom M, Lindholm L, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a cost-effectiveness analysis. *Journal of Medical Internet Research* 2017;**19**(5):e154. [SR-INCONT76798]

**Tannenbaum 2009**

Tannenbaum C, Brouillette J, Michaud J, Korner-Bitensky N, Dumoulin C, Corcos J, et al. Responsiveness and clinical utility of the geriatric self-efficacy index for urinary incontinence. *Journal of the American Geriatrics Society* 2009;**57**(3):470-5.

**Uebersax 1995**

Uebersax JS, Wyman JF, Shumaker SA, McClish DK, Fantl JA, and the Continence Program for Women Research Group. Short forms to assess life quality and symptom distress for urinary incontinence in women: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. *Neurourology and Urodynamics* 1995;**14**(2):131-9.

**Wagner 1996**

Wagner TH, Patrick DL, Bavendam TG, Martin ML, Buesching DP. Quality of life of persons with urinary incontinence: development of a new measure. *Urology* 1996;**47**(1):67-72.

**Wahl 1998**

Wahl A, Burckhardt C, Wiklund I, Hanestad BR. The Norwegian version of the Quality of Life Scale (QOLS-N). A validation and reliability study in patients suffering from psoriasis. *Scandinavian Journal of Caring Sciences* 1998;**12**(4):215-22.

**Ware 1993**

Ware JE, Snow KK, Kosinski M, Gaudek B. SF-36 Health Survey Manual and Interpretation Guide. Boston (MA): New England Medical Centre, 1993.

**WHO 2002**

World Health Organization. International Classification of Functioning, Disability and Health. Geneva: World Health Organization, 2002.

**Wilson 1999**

Wilson P, Bø K, Bourcier A, Hay-Smith J, Staskin D, Nygaard I, et al. Conservative management in women. In: Abrams P, Khoury S, Wein A editor(s). Incontinence. Plymouth: Health Publication Ltd, 1999:579-636.

**Wilson 2001**

Wilson L, Brown J, Shin G, Luc K, Subak L. Annual direct costs of urinary incontinence. *Obstetrics and Gynaecology* 2001;**98**:398-406.

**Woodley 2017**

Woodley SJ, Boyle R, Cody JD, Mørkved S, Hay-Smith EJC. Pelvic floor muscle training for prevention and treatment of urinary and faecal incontinence in antenatal and postnatal women. *Cochrane Database of Systematic Reviews* 2017, Issue 12. [DOI: 10.1002/14651858.CD007471.pub3]

**Wyman 1987**

Wyman JF, Harkins SW, Choi SC, Taylor JR, Fantl JA. Psychosocial impact of urinary incontinence in women. *Obstetrics & Gynecology* 1987;**70**(3):378-81.

**Yalcin 2003**

Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. *American Journal of Obstetrics and Gynecology* 2003;**189**(1):98-101.

**Zigmond 1983**

Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica* 1983;**67**(6):361-70.

## References to other published versions of this review

**Dumoulin 2010**

Dumoulin C, Hay-Smith J. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2010, Issue 1. [DOI: 10.1002/14651858.CD005654.pub2]

**Dumoulin 2014**

Dumoulin C, Hay Smith EJ, Mac Habée-Séguin G. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2014, Issue 5. [DOI: 10.1002/14651858.CD005654.pub3]

**Hay-Smith 2006**

Hay-Smith J, Dumoulin C. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. *Cochrane Database of Systematic Reviews* 2006, Issue 1. [DOI: 10.1002/14651858.CD005654]

* Indicates the major publication for the study

# CHARACTERISTICS OF STUDIES

## Characteristics of included studies *[ordered by study ID]*

### Aksac 2003

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design |
| | **Setting:** single centre, Turkey |
| Participants | 50 women with urodynamic SUI |
| | **Method of diagnosis:** urodynamic |
| | **Inclusion criteria:** no further criteria reported |
| | **Exclusion criteria:** no further criteria reported |
| | **Mean age (years):** PFMT 52.5 (SD 7.9); control 54.7 (SD 7.8) |
| Interventions | **Group A (n = 20):** PFMT |
| | Taught by: therapist |
| | Correct VPFMC confirmed? yes, by vaginal palpation |
| | Number VPFMC per set: 10 |
| | Number sets per day: 3 |

**Aksac 2003** *(Continued)*

| | |
|---|---|
| | Duration of hold: 5 s in the first 2 weeks, 10 s in the following weeks (3–8)<br>Duration of rest: 10 s in the first 2 weeks, 20 s in the following weeks (3–8)<br>Types of contraction e.g. submaximal, maximal: sustained<br>Duration of programme: 8 weeks<br>Body position: not reported<br>Other important information on the intervention: Home programme exercises<br>Use of digital palpation to teach VPFMC with abdominal and buttock muscle relaxation<br>Weekly clinic visits<br>**Group B (n = 10):** control, no PFMT<br>Quote: "did not undertake any exercises"<br>**Group C (n = 20):** PFM exercises via biofeedback |
| Outcomes | **Primary endpoint:** post-treatment (8 weeks)<br>**Primary outcome:** not stated<br>**Other outcomes:** pad test cure (weight gain ≤ 1 g), pad test improvement (≥ 50% reduction in pad weight), vaginal squeeze pressure, digital palpation score, incontinence frequency (4-point ordinal scale), Social Activity Index, only acquired for the intervention group.<br>On a 4-point ordinal scale (1 = urine loss once a day to 4 = urine loss once a month), median score in the PFMT group 3.5 (SD 0.5); control group 2.4 (SD 0.9)<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** not stated<br>**Funding:** not reported<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "patient was requested to choose a closed letter upon her first admission, and she was enrolled to a group in accordance with the number written in the letter"<br>No mention of sealed, opaque, consecutively numbered |
| Allocation concealment (selection bias) | Unclear risk | Quote: "patient was requested to choose a closed letter upon her first admission, and she was enrolled to a group in accordance with the number written in the letter." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the no PFMT treatment arm of the study |
| Blinding of outcome assessment (detection bias) | Unclear risk | Not clear if blinded outcome assessment |

**Aksac 2003** *(Continued)*

| | | |
|---|---|---|
| All outcomes | | |
| Incomplete outcome data (attrition bias) All outcomes | Unclear risk | No information about study completion or (n) in the results or tables |
| Selective reporting (reporting bias) | Low risk | Protocol was not available, all outcomes in study method were reported |
| Baseline comparability | Low risk | Baseline comparable for age, weight, parity, abortions, maximum birth weight, UI type |

**Asklund 2017**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br><br>**Setting:** web-based. Department of Public Health and Clinical Medicine, Unit of Research, Education and Development – Östersund, Umeå University, Sweden and Department of Public Health and Clinical Medicine, Umeå University, Sweden |
| Participants | 123 women with SUI<br><br>**Method of diagnosis:** symptoms (2-day leakage diary)<br><br>**Inclusion criteria:** women aged > 18 years, reporting at least weekly episodes of SUI for the last 6 months, with maximum voided volume ≥ 0.3 L and access to a smartphone and email.<br><br>**Exclusion criteria:** pregnancy; previous UI surgery; present or previous malignancy in the lower abdomen; impaired mobility or sensibility in the legs or lower abdomen; severe psychiatric disorders; or macroscopic haematuria, and irregular bleeding, or difficulty passing urine.<br><br>**Median age (years):** PFMT 44.8 (SD 9.7); control 44.7 (SD 9.1) |
| Interventions | **Group A (n = 62):** PFMT<br><br>Taught by: smartphone app<br><br>Correct VPFMC confirmed? basic contraction for the participant to identify the correct PFM contraction (not face-to-face)<br><br>Number VPFMC per set: progressive (from 8 to 62)<br><br>Number sets per day: 3<br><br>Duration of hold: 2–59 s<br><br>Duration of rest: 2–59 s<br><br>Type(s) of contraction e.g. submaximal, maximal: exercises to find the right muscle, train strength, endurance, quick contraction and knack<br><br>Duration of programme: 3 months<br><br>Position: standing, lifting, walking<br><br>Other important information on the intervention:<br><br>Application contained information about SUI, PFMT exercises at different levels (6 basic and 6 advanced) with graphic support, and functions for statistics and reminders + lifestyle, overweight, smoking physical activity and drinking. There was no face-to-face contact with the participants during study. |

**Asklund 2017** *(Continued)*

| | |
|---|---|
| | **Group B (n = 61):** control<br><br>Treatment postponed, receiving application after 3 months' follow-up |
| Outcomes | **Primary endpoint:** post-treatment (3 months)<br><br>Outcome measured during the intervention: performing PFMT (yes or no), weekly and daily for both groups<br><br>**Primary outcome:** symptom severity (International Consultation on ICIQ-UI Short Form); condition-specific QoL (ICIQ-LUTSqol)<br><br>**Secondary outcome:** PGI-I; IEF; usage of incontinence aids; participant satisfaction<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Subsequent publications:**<br><br>Lindh A (International Urogynecology Journal 2016;27(12):1857-65) documented predictors of success after 1 year in the PFMT group. Participants with successful short-term results were more likely to succeed in the corresponding outcome at 1 year than those without successful short-term results (adjusted OR): PGI 5.15, 95% CI 2.40 to 11.03; ICIQ-UI Short Form 6.85, 95% CI 2.83 to 16.58; and sufficient treatment 3.78, 95% CI 1.58 to 9.08. Increasing age predicted success in PGI-I (adjusted OR 1.06, 95% CI 1.02 to 1.10) and sufficient treatment (adjusted OR 1.08, 95% CI 1.03 to 1.13). Compared with not training regularly, regular PFMT at 1 year predicted success for PGI (adjusted OR 2.32, 95% CI 1.04 to 5.20) and sufficient treatment (adjusted OR 2.99, 95% CI 1.23 to 7.27)<br>.<br><br>Hoffman (Acta Obstetricia et Gynecologica Scandinavica 2017;96(10):1180-7) presented a 2-year follow-up of the RCT for the PFMT group only. 75.4% participated in 3-month and 2-year follow-ups. Baseline data did not differ between responders and non-responders at follow-up. After 2 years, mean decrease in ICIQ-UI Short Form was 3.1 (95% CI 2.0 to 4.2) and ICIQ-LUTqol was 4.0 (95% CI 2.1 to 5.9). Of the 46 women, 8.7% rated themselves as very much better, 19.6% as much better and 34.8% as a little better.<br><br>Sjostrom M (Journal of Medical Internet Research 2017;19(5):e154) documented cost-effectiveness through annual costs per group (app group Euro 447.0 and control group Euro 482.4**),** annual gains in quality-adjusted life years per group (app group 0.0101 and control group 0.0016), and the extra cost per quality-adjusted life year compared to control group ranged from Euro –2425.7 to Euro 14,870.6, which indicated greater gains in quality-adjusted life years at similar or slightly higher cost.<br><br>**Dropouts and withdrawal:** total 2/123 (PFMT 1/62; control 1/61). Intention-to-treat analysis.<br><br>**Funding:** "This study was supported by grants from the Swedish Research Council for Health, Working Life and Welfare, the Jämtland County Council, the Västerbotten County Council, and VisareNorr, Northern County Councils, Sweden."<br><br>**Conflicts of interest:** None. Quote: "None of the researchers have any financial interests in the product." |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Generated by independent administrator |
| Allocation concealment (selection bias) | Low risk | Randomisation was performed with allocation concealment using sequentially numbered, opaque, sealed envelopes without blinding. |

**Asklund 2017** *(Continued)*

| | | |
|---|---|---|
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the postponed treatment arm of the study. |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | No assessor, participants completed the questionnaires themselves. |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Total of 2 (2%), reasons unknown.<br>By group: PFMT = 1 (2%); control = 1 (2%)<br>+ incomplete outcome for the app group (–2 to –3) in 3 outcomes<br>Intention-to-treat analysis used |
| Selective reporting (reporting bias) | Low risk | Trial protocol available through a trial registry, all outcomes in the protocol reported in the trial. |
| Baseline comparability | Low risk | Groups were comparable at baseline. |

**Bertotto 2017**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design<br>**Setting:** single centre (outpatient clinic of Centro Universitario La Salle, Canoas, RS), Brazil |
| Participants | 49 postmenopausal women<br>**Method of diagnosis:** symptoms (ICIQ-UI Short Form: loss of urine on exertion)<br>**Inclusion criteria:** women in postmenopausal status, aged 50–65 years, complaint of urine loss on exertion (detected by ICIQ-UI Short Form questionnaire), provision of written informed consent<br>**Exclusion criteria:** presence of urinary tract infection, failure to understand PFM contraction, cognitive alterations, collagen- or muscle-related diseases, neurological abnormalities<br>**Mean age (years):** PFMT 59.3 (SD 4.9); control 57.1 (SD 5.3) |
| Interventions | **Group A (n = 16):** PFMT<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? yes, by digital palpation<br>Number VPFMC per set: progressive (from 24 to 70)<br>Number sets per day: progressive (1–3)<br>Duration of hold: 6–10 s, depending on the exercise<br>Duration of rest: 4–10 s, depending on the exercise<br>Type(s) of contraction, e.g. submaximal, maximal: sustained, phasic lasting 2 s, phasic lasting 3–5 s, guides-imagery training to generate or enhance precontraction to increased abdominal pressure<br>Duration of programme: 4 weeks<br>Position: supine, seated, standing as participant improved |

**Bertotto 2017** *(Continued)*

| | |
|---|---|
| | **Group B (n = 16):** control<br><br>Received no treatment in the intervening period<br><br>**Group C (n = 17):** PFM exercises + biofeedback |
| Outcomes | **Primary endpoint:** post-treatment (4 weeks)<br><br>**Primary outcome:** not specified which are primary and secondary<br><br>Precontraction, initial EMG baseline (μv), final EMG baseline (μv), duration of endurance contraction (s), maximum voluntary contraction (μv), ICIQ-UI Short Form<br><br>**Outcome measured during the intervention:** none reported<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** total 3/32 (PFMT 1/16; control 2/16). No intention-to-treat analysis<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Women who met the inclusion criteria were randomised across 3 groups. Randomisation was performed using envelopes containing the letters A, B and C, where each letter corresponded to a specific group to which the participant would be allocated, by order of presentation to the study facility |
| Allocation concealment (selection bias) | Low risk | Randomisation was performed using envelopes containing the letters A, B and C, where each letter corresponded to a specific group to which the participant would be allocated, by order of presentation to the study facility<br><br>This allocation was performed by a blinded, independent researcher not otherwise involved in the study |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the control arm |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | The study was conducted by 2 investigators (1 and 2). Assessment of groups before and after intervention was performed by investigator 1, while training was performed by investigator 2. Both had been previously trained<br><br>No mention if the assessor was blinded to the intervention |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Total of 3 (9.4%), reasons unknown<br><br>By group: PFMT = 1 (6.3%); control = 2 (12.5%)<br><br>No intention-to-treat analysis used |
| Selective reporting (reporting bias) | Low risk | Trial protocol available through a trial registry, all outcomes in the protocol reported in the trial |

**Bertotto 2017** *(Continued)*

| | | |
|---|---|---|
| Baseline comparability | Low risk | Groups were comparable at baseline |

**Beuttenmuller 2010**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT<br>**Setting:** single centre; the Rehabilitation Unit of Pelvic Floor Disorders in Fortaleza-Ceara, Brazil |
| Participants | 75 women with SUI<br>**Method of diagnosis:** not reported. Quote: "women with a diagnosis of SUI"<br>**Inclusion criteria:** not reported<br>**Exclusion criteria:** not reported<br>**Mean age (years):** PFMT 49.96 (SD 5.26); control 44.82 (SD 4.88) |
| Interventions | **Group A (n = 25):** PFMT<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? not reported but assessed by the evaluator prior to treatment<br>Number VPFMC per set: 8<br>Number sets per day: not reported<br>Duration of hold: 5 s<br>Duration of rest: not reported<br>Type(s) of contraction e.g. submaximal, maximal?: long and short contraction with the participant in supine lying position with knee bent, sitting in the chair or on the gym ball, on all fours, standing<br>Duration of programme: 20 min (in groups of 4) twice per week for 6 weeks except during menstrual periods or due to other complications<br>Number and type of contact with health professional(s): twice per week<br>Other information: kinesitherapy was accomplished through standing or sitting exercises using a Swiss ball of varying size, according to the height and weight of the participant. Proprioceptive exercises such as hopping on a ball, moves to raise the pelvis (anteversion, retroversion, lateralisation and circumduction) were used. Additionally, exercises were used to contract the PFM to the original position, working the 2 fibre types I and II by performing contract-relax perineal exercises and hold-relax training, respectively, up to 6 s.<br>**Group B (n = 25):** control. Quote: "no physical therapy at that time"<br>**Group C (n = 17):** PFM exercises with electrical stimulation |
| Outcomes | **Primary endpoint:** post-treatment (6 weeks)<br>**Primary outcome:** not specified which are primary and secondary<br>KHQ, PFM 1 finger intravaginal evaluation using the Oxford scale, intravaginal pressure perineometry<br>**Outcome measured during the intervention:** none reported<br>**Any outcomes measured but not reported:** no |
| Notes | **Dropouts and withdrawal:** not clear |

**Beuttenmuller 2010** *(Continued)*

**Funding:** none. Quote: "No funds were received in support of this study."

**Conflicts of interest:** none. Quote: "No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript."

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "randomly divided in 3 groups" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not clear if outcome assessment was blinded |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Not clear if there was attrition |
| Selective reporting (reporting bias) | Low risk | Protocol not available, all outcomes in study method were reported. |
| Baseline comparability | Low risk | Groups comparable at baseline for age and BMI |

**Bidmead 2002**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design (after treatment period, control participants crossed over into group C (PFMT with active electrical stimulation))<br>**Setting:** single centre, UK |
| Participants | 170 women with urodynamic SUI<br>**Inclusion criteria:** new diagnosis of SUI or no treatment for SUI in previous 6 months<br>**Exclusion criteria:** no further criteria reported<br>**Mean age (years):** PFMT 46.2 (SD 8.5); control 47.5 (SD 11.5) |
| Interventions | **Group A (n = 40):** PFMT<br>Conventional PFMT supervised by physiotherapist<br>Individually tailored lifestyle advice.<br>5 clinic visits in 14 weeks (weeks 1, 3, 6, 10 and 14)<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? not reported |

**Bidmead 2002** *(Continued)*

| | |
|---|---|
| | Number VPFMC per set: not reported<br>Number sets per day: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Type(s) of contraction, e.g. submaximal, maximal: not reported<br>Duration of programme: 14 weeks<br>Position: not reported<br>**Group B (n = 20):** control<br>No treatment for 14 weeks<br>Thereafter crossed over into group C (PFMT with active electrical stimulation)<br>**Group C (n = 88):** PFM exercises with active electrical stimulation<br>**Group D (n = 42):** PFM exercises with sham electrical stimulation |
| Outcomes | **Primary endpoint:** post-treatment (14 weeks)<br>**Primary outcome:** not specified which are primary and secondary. Pad test, KHQ<br>**Outcome measured during the intervention:** none reported<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** total 17/60 (PFMT 10/40; control 7/20). Primary analysis by intention-to-treat<br>**Funding:** not reported<br>**Conflicts of interest:** not reported |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "randomised" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate random allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the cross-over control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Similar in each of the 4 groups at around 25%<br>Primary analysis by intention-to-treat |

**Bidmead 2002** *(Continued)*

| | | |
|---|---|---|
| Selective reporting (reporting bias) | Unclear risk | Reporting on all outcomes but not with data for all of them (abstract available only) |
| Baseline comparability | Low risk | Groups comparable at baseline for age, severity, severity of genuine stress incontinence on urodynamics |

**Burgio 1998**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design<br><br>Stratified by type (UUI, MUI) and severity of incontinence (number of leakage episodes)<br><br>**Setting:** single centre, USA |
| Participants | 197 women with UUI, with or without urodynamic SUI<br><br>**Inclusion criteria:** community-dwelling women aged ≥ 55 years, ≥ 2 urge accidents per week, UUI-predominant pattern<br><br>**Exclusion criteria:** continual leakage, uterine prolapse past introitus, unstable angina, decompensated heart failure, history of malignant arrhythmias, impaired mental status (MMSE < 20)<br><br>**Mean age (years):** PFMT 67.3 (SD 7.6); control 67.6 (SD 7.6)<br><br>Mean duration of symptoms (years): PFMT 9.4 (SD 10.8); control 12.7 (SD 15.9)<br><br>> 10 leakage episodes per week: PFMT 52%; control 54%<br><br>**Diagnosis:** 96 UUI only (49%); 101 MUI (51%) |
| Interventions | **Group A (n = 65):** PFMT<br><br>Taught by: nurse<br><br>Correct VPFMC confirmed? anorectal biofeedback<br><br>Number VPFMC per set: 15<br><br>Number sets per day: 3<br><br>Duration of hold: up to 10 s<br><br>Duration of rest: up to 10 s<br><br>Quote: "duration of individual contraction and relaxation was based on the ability demonstrated by each patient in the biofeedback sessions and gradually increased across sessions to a maximum of 10 seconds each."<br><br>Type(s) of contraction e.g. submaximal, maximal: sustained maximal contractions and quick contractions<br><br>Duration of programme: 8 weeks<br><br>Position: various. Quote: "patients were advised to practice in various positions, including lying, sitting, and standing."<br><br>Other important information on the intervention:<br><br>Use of anorectal biofeedback to teach VPFMC with abdominal muscle relaxation. Response to urge (pause, sit, relax, repeated VPFMC to suppress urge). Use of bladder-sphincter biofeedback at 3rd visit for women with < 50% reduction in leakage episodes to teach VPFMC against increasing fluid volume and urge. Fortnightly clinic visit with nurse practitioner, 8 weeks |

**Burgio 1998** *(Continued)*

| | |
|---|---|
| | **Group B (n = 65):** control<br><br>Placebo drug, 3 times a day, for 8 weeks. Capsule contained riboflavin phosphate 500 mg marker. Fortnightly clinic visit with nurse practitioner<br><br>**Group C (n = 67):** drug treatment |
| Outcomes | **Primary endpoint:** post-treatment (10 weeks)<br><br>**Primary outcome:** change in leakage frequency (2-week urinary diary)<br><br>**Secondary outcomes:** Hopkins Symptom checklist for psychological distress, self-report (worse to much better), satisfaction with progress (not at all to completely), perceived improvement (none or 0% to dry or 100%), willingness to continue PFMT, desire for other treatment, leakage episodes (2-week urinary diary), cystometry (for 105/197), adverse events<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** total 16/130 (PFMT 4/65; control 12/65). Intention-to-treat for primary outcome, most recent urinary diary data carried forward<br><br>**Funding:** "This research was supported by grants AG 08010 and KO4 00431 from the National Institute on Aging, National Institutes of Health, Bethesda, Md."<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "within each stratum, randomisation was performed with computer-generated random numbers using a block size of 6 to avoid inequity in group size." |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate allocation concealment<br><br>Quote: "within each stratum, randomisation was performed with computer-generated random numbers using a block size of 6 to avoid inequity in group size." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Control group was placebo-blinded to the medication treatment arm but not to the PFMT group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Attrition rate per group and reasons given: not thought to be due to intervention except for 1 participant in the placebo groups<br><br>Primary analysis by intention-to-treat |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |

**Burgio 1998** *(Continued)*

| | | |
|---|---|---|
| Baseline comparability | Low risk | Quote: "Before treatment the groups were comparable on all key parameters except that subject in behavioral treatment had more children, were less likely to have a high school education and more likely to have a rectocele." |

**Burns 1993**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design<br><br>**Setting:** single centre, USA |
| Participants | 135 women with urodynamic SUI, with or without DO<br><br>**Inclusion criteria:** women with SUI or MUI, aged ≥ 55 years, ≥ 3 leakage episodes per week, demonstrates leakage with stress manoeuvres during physical examination, MMSE > 23, absence of glycosuria or pyuria, postvoid residual volume < 50 mL, maximum uroflow > 15 mL/s.<br><br>**Exclusion criteria:** no additional criteria reported<br><br>**Mean age (years):** PFMT 63 (SD 6); control 63 (SD 5)<br><br>Mean leakage episodes per 24 hours: PFMT 2.6 (SD 2.1); control 2.6 (SD 2.6)<br><br>**Diagnosis:** 135 urodynamic SUI |
| Interventions | **Group A (n = 43, after dropouts):** PFMT<br><br>Taught by: nurse<br><br>Correct VPFMC confirmed? biofeedback<br><br>Number VPFMC per set: 20 (10 quick and 10 sustained) progressive (increased by 10 per set to a total of 200/day)<br><br>Number sets per day: 4<br><br>Duration of hold: 3 s (quick) to 10 s (sustained)<br><br>Duration of rest: not reported<br><br>Type(s) of contraction e.g. submaximal, maximal: quick and sustained<br><br>Duration of programme: 8 weeks<br><br>Position: not reported<br><br>Other important information on the intervention:<br><br>Booklet explaining anatomy, PFMT, and completion of exercise and urinary diaries. Videotape describing exercise protocol. Weekly exercise reminder cards mailed between visits. Weekly clinic visits with nurse<br><br>**Group B (n = 40, after dropouts):** control, no treatment<br><br>**Group C (n = 40, after dropouts):** PFMT with biofeedback |
| Outcomes | **Primary endpoint:** post-treatment (8 weeks)<br><br>**Primary outcome:** leakage episodes (2-week urinary diary)<br><br>**Secondary outcomes:** incontinence severity (based on number of leakage episodes from diary), PFM EMG, cystometry |

**Burns 1993** *(Continued)*

| | |
|---|---|
| | **Follow-up after primary endpoint:** longer-term follow-up at 12 weeks and 6 months |
| Notes | **Dropouts and withdrawal:** 12 total (10/135 dropout and 2/135 excluded from analysis (no urinary diary); group not specified)<br><br>**Funding:** "This research was funded by a cooperative agreement (UOI AG05260) from the National Institute on Aging and the National Center for Nursing Research."<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "randomised blocking was employed to balance the number of subjects in each group." |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the no treatment control arm of the study |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | 10/135 dropped out or withdrawn, 2 did not have bladder diary data so were excluded from analysis |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Sociodemographic comparable |

**Bø 1999**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design<br><br>Stratified by severity of leakage on pad test<br><br>A priori power calculation<br><br>**Setting:** 5 centres, Norway |
| Participants | 122 women with urodynamic SUI<br><br>**Inclusion criteria:** women with a history of SUI, waiting for surgery or recruited through advertising, > 4 g leakage on pad test with standardised bladder volume<br><br>**Exclusion criteria:** other types of incontinence, DO on urodynamics, residual urine > 50 mL, maximum uroflow < 15 mL/s, previous surgery for urodynamic SUI, neurological or psychiatric disease, ongoing urinary tract infection, other disease that could interfere with participation, use of concomitant treatments during trial, inability to understand instructions given in Norwegian |

**Bø 1999** *(Continued)*

| | |
|---|---|
| | **Mean age (years):** PFMT 49.6 (SD 10.0); control 51.7 (SD 8.8)<br><br>Mean duration of symptoms (years): PFMT 10.2 (SD 7.7); control 9.9 (SD 7.8)<br><br>Mean leakage episodes per 24 hours: PFMT 0.9 (SD 0.6); control 1.0 (SD 1.0)<br><br>**Diagnosis:** 122 urodynamic SUI (100%) |
| Interventions | **Group A (n = 29):** PFMT<br><br>Taught by: physiotherapist<br><br>Correct VPFMC confirmed? yes, by vaginal palpation<br><br>Number VPFMC per set: 8–12 sustained contractions + 3–4 fast contractions<br><br>Number sets per day: 3<br><br>Duration of hold: 6–8 (for the sustained contraction) s<br><br>Duration of rest: 6 s<br><br>Type(s) of contraction, e.g. submaximal, maximal: high intensity (close to maximum) fast and sustained contractions<br><br>Duration of programme: 6 months<br><br>Position: lying, standing, kneeling and sitting positions with legs apart<br><br>Other important information on the intervention:<br><br>Explanation of anatomy, physiology and continence mechanism by physiotherapist. Audiotape of home training programme. Weekly 45-min exercise class with PFMT in a variety of body positions, and back, abdominal, buttock and thigh muscle exercises. Monthly clinic visit with physiotherapist<br><br>**Group B (n = 32):** control<br><br>Explanation of anatomy, physiology and continence mechanism. Correct VPFMC confirmed by palpation. No clinic visits. Offered instruction in use of the Continence Guard (14 accepted)<br><br>**Group C (n = 32):** electrical stimulation<br><br>**Group D (n = 29):** vaginal cones |
| Outcomes | **Primary endpoint:** post-treatment evaluation at 6 months<br><br>**Primary outcomes:** 60 s pad test with standardised (200 mL) bladder volume (30 s of running on the spot + 30 s of jumping with legs in subsequent adduction and abduction (jumping jacks, rate of 132 beats/min)), self-report (very problematic to unproblematic)<br><br>**Secondary outcomes:** The Norwegian version of the Quality of Life Scale (QoLS-N), Bristol Female Lower Urinary Tract Symptoms Questionnaire, Leakage Index, Social Activity Index, leakage episodes (3-day urinary diary), 24-hour pad test, vaginal squeeze pressure, adverse events<br><br>**Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts and withdrawal:** total 6/61 (PFMT 4/29; controls 2/32). Intention-to-treat: baseline values used for losses to follow-up<br><br>**Funding:** "Norwegian Fund for Postgraduate studies in Physiotherapy and Norwegian Research Council. Coloplast AS provided the continence guards and Vitacon AS provided the electrical stimulators and cones. They also gave financial support to seminars for the research group."<br><br>**Conflicts of interest:** none declared |

**Bø 1999** *(Continued)*

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "computer generated random number" |
| Allocation concealment (selection bias) | Low risk | Publication stated "random." Contact with author confirmed random number generation, and sealed opaque envelopes |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the no treatment control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Quote: "physicians evaluating the effect of the treatment were also blind to allocation of treatment." |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Attrition details: 3 could not complete the study (asthma, change of work, death in the family), 2 were excluded because they used other treatment during the trial. Dropout: 2 from PFMT (8%) (motivation, travel time) and 0 from control group (0%)<br><br>Secondary analysis by intention-to-treat |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Groups were comparable at baseline |

**Carneiro 2010**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT<br><br>**Setting:** single centre, Cafisio Physical Therapy Clinic |
| Participants | 50 women aged 30–55 years with SUI<br><br>**Method of diagnosis:** urodynamic<br><br>**Inclusion criteria:** women referred by urologists and gynaecologists with urodynamic diagnosis of SUI due to bladder neck hypermobility or PDS ≥ 90 $cmH_2O$<br><br>**Exclusion criteria:** SUI due to intrinsic insufficiency (PDS) < 60 $cmH_2O$, prior surgical correction of SUI and genital prolapse of any grade in physical examination<br><br>**Mean age (years):** PFMT 49.24 (SD 7.37); control 45.25 (SD 6.60) |
| Interventions | **Group A (n = 25):** PFMT<br><br>Taught by: physical therapist<br><br>Correct VPFMC confirmed? yes and maximum voluntary contraction was verified by initial assessment, individually for each women<br><br>Number VPFMC per set: 8–12 repetitions of 5 perineal exercises |

**Carneiro 2010** *(Continued)*

| | |
|---|---|
| | Number sets per day: once<br><br>Duration of hold: 6–10 s<br><br>Duration of rest: not mentioned<br><br>Type(s) of contraction e.g. submaximal, maximal: not reported<br><br>Duration of programme: 30 min, twice per week for 8 consecutive weeks<br><br>Number and type of contact with health professional(s): twice per week<br><br>Other information: verbal information about the PFM function, visualisation of PF component with anatomical figures<br><br>Five minutes of proprioception sitting on a 75-cm diameter therapeutic ball. During that time, participants were asked to make lateral movements of the pelvis, pelvic anteversion movements, short jumps and figure of 8 movement with the pelvis<br><br>**Group B (n = 25):** control<br><br>Quote: "The control group carried out no activity during the 8 weeks, as they were on the waiting list." |
| Outcomes | **Primary endpoint:** post-treatment evaluation at 8 weeks<br><br>**Primary outcome:** not stated<br><br>**Other outcomes:** ultrasound examination, surface EMG with an intravaginal probe, PFM bidigital muscle strength test, KHQ<br><br>**Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts and withdrawal:** not stated<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "Using a simple random sampling" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the no treatment control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not clear if outcome assessment was blinded |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Attrition not reported |

**Carneiro 2010** *(Continued)*

| | | |
|---|---|---|
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Quote: "Groups comparable for age, vaginal delivery, caesarian delivery and time with UI."<br><br>Quote: "Time with UI was almost significantly different between the two group with the Group A having had UI for a longer time." |

**Castro 2008**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design<br><br>**Setting:** single centre? Sao Paulo, Brazil |
| Participants | 118 women with urodynamic SUI without DO<br><br>**Inclusion criteria:** women with urodynamic SUI, no DO, a positive cough test, > 3 g leakage measured on pad test with standardised bladder volume (200 mL); mean of 3 episodes of UI per week<br><br>**Exclusion criteria:** chronic degenerative disease that would affect muscular or nerve tissues, advanced genital prolapse, pregnancy, active or recurrent UTI, vulvovaginitis, atrophic vaginitis, continence surgery within 1 year, people with pacemaker, Valsalva leak point pressure < 60 $mmH_2O$ in sitting with 250 mL in bladder or UCP < 20 $cmH_2O$ in sitting position at maximal cystometric capacity<br><br>**Mean age (years):** PFMT 56.2 (SD 12.5); control 52.6 (SD 11.2)<br><br>Leakage episodes in 7 days: PFMT 10.3 (SD 10.1); control 10.5 (SD 7.0)<br><br>**Mean BMI ($kg/m^2$):** PFMT 25.9 (SD 5.0); control 26.9 (SD 5.1) |
| Interventions | **Group A (n = 31):** PFMT<br><br>Taught by: physiotherapist<br><br>Correct VPFMC confirmed? yes, by vaginal palpation<br><br>Number VPFMC per set: 60<br><br>Number sets per day: 3 per week<br><br>Duration of hold: 1 s, 2 s, 5 s and 10 s<br><br>Duration of rest: same as contraction (1 s, 2 s, 5 s and 10 s)<br><br>Type(s) of contraction e.g. submaximal, maximal: sustained for 1–10 s<br><br>Duration of programme: 6 months<br><br>Body position: not reported<br><br>Other important information on the intervention: Home programme exercises (3 × 45-min exercises classes per week (including PFMT) for 6 months with supervision by physiotherapist)<br><br>**Group B (n = 30):** control<br><br>No visit with therapist but motivational telephone calls once per month<br><br>**Group C (n = 30):** electrical stimulation<br><br>**Group D (n = 27):** vaginal cones |

**Castro 2008** *(Continued)*

| | |
|---|---|
| Outcomes | **Primary endpoint:** post-treatment (6 months)<br><br>**Primary outcomes:** objective cure of SUI based on a negative pad test with a standardised bladder volume (< 2 g in weight)<br><br>**Secondary outcomes:** I-QOL, voiding diary (number of leakage in 7 days), PFM digital evaluation using Oxford scale, urodynamics evaluation, subjective cure "satisfied" or "dissatisfied," adverse events<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts and withdrawal:** total 11/61 (2/31 + 3/31 excluded from PFMT, 4/30 + 2/30 excluded from controls). No Intention-to-treat analysis<br><br>Quote: "We analyzed the data only for those women who completed the study."<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Computer-generated random numbers |
| Allocation concealment (selection bias) | Low risk | Quote: "Once enrolled by a physician investigator, subjects were assigned to four distinct groups: pelvic floor exercises, electrical stimulation, vaginal cones, or untreated controls. The division of the four groups was undertaken by using computer-generated random numbers prepared by the Biostatistics Center of the Federal University of São Paulo." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Dropout/excluded<br><br>PFMT (n = 5): 2 lack of improvement + 1 pregnancy + ?<br><br>Control (n = 6): 2 lack of improvement + 3 change of city + ? |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Quote: "there were no significant differences between the groups in any of demographics, clinical characteristics or outcome measurements." |

**Celiker Tosun 2015**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design |

**Celiker Tosun 2015** *(Continued)*

| | |
|---|---|
| | **Setting:** jointly conducted by the Ege University Faculty of Medicine, Department of Obstetrics and Gynecology and the Dokuz Eylul University, School of Physical Therapy and Rehabilitation |
| Participants | 121 women with UI<br><br>**Method of diagnosis:** sign (urodynamic)<br><br>**Inclusion criteria:** women with SUI and MUI diagnosed by a urogynaecologist using analysis of urodynamics<br><br>**Exclusion criteria:** pregnancy, history of spinal surgery, spinal or pelvic fracture, urinary tract infection, vaginal infection, known neurologic disorders, respiratory disease and menstruation at the time of assessment, PFMT at physiotherapy within the last 2 years and 'zero-one' PFM strength (digital palpation)<br><br>**Mean age (years):** PFMT 51.7 (SD 10.3); control 52.5 (SD 9.1) |
| Interventions | **Group A (n = 58): PFMT**<br><br>Taught by: physiotherapist<br><br>Correct VPFMC confirmed? Yes, by digital palpation<br><br>Number VPFMC per set: 12<br><br>Number sets per day: 3<br><br>Duration of hold: according to Laycock PERFECT scheme<br><br>Duration of rest: according to Laycock PERFECT scheme<br><br>Type(s) of contraction e.g. submaximal, maximal: progressive training with fast, slow and sustained contractions + knack<br><br>Duration of programme: 30 min. 3 times per week for the first 2 weeks followed by weekly visits + additional programme for those who did not reach the PFM strength goal (grade 5 on Oxford scale)<br><br>Position: not reported<br><br>**Group B (n = 63):** control, no treatment waiting list |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br><br>**Primary outcome:** IIQ-7, UDI-6, urgent voiding urinary, daytime urinary frequency, UI frequency, nighttime urinary frequency, 1-hour pad test, stop test<br><br>**Secondary outcome:** PFM strength by digital palpation (PERFECT), morphometry by transabdominal ultrasound (Ultrasonix-ES500, Canada), intravaginal pressure by manometry (Peritron)<br><br>**Outcome measured during the intervention:** PERFECT scheme not reported<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** for a subgroup after 8–12 weeks of additional training to achieve the PFM strength goal (grade 5 on Oxford scale) |
| Notes | **Dropouts and withdrawal:** not stated<br><br>**Funding:** "This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors."<br><br>**Conflicts of interest:** none declared |
| ***Risk of bias*** | |

**Celiker Tosun 2015** *(Continued)*

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Computer-generated random number table |
| Allocation concealment (selection bias) | Low risk | Prelabelled sealed envelope |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the control arm |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | The investigators were not aware of the results of the other assessments<br>1 physiotherapist performed measurements and the other performed PFMT |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Total of 9/130 (7%)<br>By group: PFT 7/65 (11%); control 2/65 (3%)<br>Unbalanced but low attrition rate in each group<br>No intention-to-treat analysis used |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section, no trial registry found |
| Baseline comparability | Unclear risk | Yes for age, BMI, waist/hip circumference, duration of symptoms, number of pregnancy, heaviest new born weight at delivery<br>No for IIQ-7 and time frequency<br>(PFMT group had lower IIQ-7 scores and higher night-time urinary frequency before the intervention) |

**Diokno 2010**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT<br>**Setting**: 4 Michigan counties |
| Participants | 44 adult ambulatory women with UI<br>**Method of diagnosis:** symptoms of UI on the MESA questionnaire<br>**Inclusion criteria:** MESA score showing UI. Quote: "Previously failed anti-incontinence surgery was not considered exclusion."<br>**Exclusion criteria:** currently under treatment for UI; history of bladder cancer, stroke, multiple sclerosis, Parkinson's, epilepsy, spinal cord tumour or trauma; pregnancy; MESA ≥ 72% for urge score or MESA ≥ 70% for stress score, in addition to urge percentage higher than stress percentage (to eliminate women with total incontinence and women with UUI predominant symptoms, respectively)<br>**Mean age (years):** PFMT 60.6 (SD 14.4); control 52.2 (SD 12.6) |
| Interventions | **Group A (n = 23):** PFMT<br>Taught by: urology nurse |

**Diokno 2010** *(Continued)*

| | |
|---|---|
| | Correct VPFMC confirmed? Yes, vaginal examination to test for PFM strength were performed by 2 nurses<br><br>Number VPFMC per set: 25 contractions in lying and other positions (5 short contractions (quick squeezes) and 20 long contractions (hold up to 6 s) + knack when needed (sneezing)<br><br>Number sets per day: twice per day<br><br>Duration of rest: not reported<br><br>Type(s) of contraction e.g. submaximal, maximal: maximal<br><br>Duration of programme: 1 teaching session, 1 follow-up session and daily exercises with an audiotape of PFMT<br><br>Number and type of contact with health professional(s): once after 2–4 weeks with vaginal examination if needed and written test on new knowledge acquired<br><br>Other important information:<br><br>Bladder training tips if needed: progressive voiding schedule based on participant's diary done before attending the class, interval increased by 15–30 min, use pelvic muscle contraction and distraction to inhibit detrusor. Goal: voiding interval of 3.5–4 hours while awake. This was not applicable if they already have the 3.5- to 4-hour interval at baseline.<br><br>Quote: "2-h power point presentation lecture in groups by two trained urology nurses. Paper handouts were distributed."<br><br>**Group B (n = 21):** control<br><br>No information given on behavioural intervention at any time |
| Outcomes | **Primary endpoint:** post-treatment (8 weeks)<br><br>**Primary outcome:** not stated<br><br>**Other outcomes:** improvement, as measured by reduction of severity level on a 3-point scale (severe to moderate or mild and moderate to mild), or 'no-improvement' for those who stayed the same or worsened, voiding frequency/intervoid interval, continence status with pad testing (grams), cough test leak diameter (in centimetres), stress test (percentage positive) and PFM strength with digital score (pressure, displacement, duration)<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** not stated<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "randomisation was performed in groups of five using the SAS system." |
| Allocation concealment (selection bias) | Unclear risk | Adequate allocation concealment |

**Diokno 2010** *(Continued)*

| | | |
|---|---|---|
| Blinding of participants and personnel (performance bias) All outcomes | Unclear risk | Blinding of the participants was not feasible for the control arm |
| Blinding of outcome assessment (detection bias) All outcomes | Low risk | Vaginal examinations to test PFM strength and collection of bladder diary and 24-hour pad test were performed by 2 nurses other than the lecturers |
| Incomplete outcome data (attrition bias) All outcomes | Low risk | Total attrition: 1 could not contract and did not get randomised, so 44/45 women participated to randomisation<br>Group A: 0/23 (0%)<br>Group B: 3/21 (14%) reason: had incomplete data |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | High risk | Quote: "The only demographic statistically significant difference between the two groups was in age."<br>Women in the treatment group were older |

**Ferreira 2014**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br>**Setting:** Familicao Athletic Volleyball Club |
| Participants | 32 women with SUI<br>**Method of diagnosis:** symptoms and sings of SUI (bladder diary + pad test)<br>**Inclusion criteria:** women aged 13–30 years, nulliparous, with symptoms of stress UI (7-day bladder diary), with urinary leakage > 1 g (pad test: 15 min of volleyball practice)<br>**Exclusion criteria:** treatment for < 6 months, sport practice for < 2 years, currently or repeated urinary infections, 18 > BMI > 25<br>**Mean age (years):** PFMT 19.4 (SD 3.24); control 19.1 (SD 2.11) |
| Interventions | **Group A (n = 16):** PFMT<br>Taught by: not reported<br>Correct VPFMC confirmed? unclear<br>Number VPFMC per set: 30 sustained contractions + 4 × 30 fast contractions<br>Number sets per day: 1<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Type(s) of contraction e.g. submaximal, maximal: not reported<br>Duration of programme: 3 months<br>Position: not reported |

**Ferreira 2014** *(Continued)*

| | |
|---|---|
| | Other important information on the intervention: none<br><br>**Group B (n = 16):** control, had access to an educational pamphlet |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br><br>**Primary outcome:** questionnaire (not specified, only pretreatment data reported), frequency of incontinence, pad test<br><br>**Other outcomes:** none<br><br>**Outcome measured during the intervention:** none reported<br><br>**Any outcomes measured but not reported:** questionnaire not specified<br><br>**Follow-up after primary endpoint:** no |
| Notes | Adherence < 50% was reported as an exclusion criteria; however, no measures of adherence were specified.<br><br>**Dropouts and withdrawal:** no losses<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Lottery |
| Allocation concealment (selection bias) | Unclear risk | Folded pieces of paper were placed in a common box |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the control arm, although both treatment groups were unaware of the other treatment arm |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not reported |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | No losses |
| Selective reporting (reporting bias) | High risk | No reporting of the results of an outcome (questionnaire) described in the methods |
| Baseline comparability | Low risk | Groups were comparable at baseline |

**Firra 2013**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design |

**Firra 2013** *(Continued)*

| | |
|---|---|
| | **Setting:** Texas Woman's University, USA |
| Participants | 64 women with UI (38 SUI + 26 UUI)<br><br>**Method of diagnosis:** signs and symptoms<br><br>**Inclusion criteria:** parous or nulliparous women, aged ≥ 21 years, with a diagnosis of stress or UUI, with manual dexterity to dial the Liberty Electrical Stimulation Unit (Utah Medical, Midvale, UT), able to understand and follow instructions in English, women taking hormone replacement therapy had to agree to maintain the same oestrogen intake over the course of the study, women not taking hormones were asked not to start an oestrogen regimen during the study<br><br>**Exclusion criteria:** scored 0 on the Oxford scale for PFM strength, had denervation injury to the sphincters, had undergone anti-incontinence surgery, had vaginal stenosis to the extent that the middle finger could not be inserted into the vagina, BMI > 50, women with stage III or IV prolapse or current bladder infection (excluded until cured), pregnant, with neurological conditions, taking any potentially confounding prescription drug or nutraceutical (e.g. tolterodine, oxybutynin, antidepressants, antihistamines, black cohosh), or those with < 3 UI episodes in 3 days<br><br>**Mean age (years):** SUI PFMT 63.6 (SD 13.3); SUI control 48.2 (SD 16.2); UUI PFMT 66.5 (SD 12.4); UUI control 63.0 (SD 14.5) |
| Interventions | **Group A:** PFMT; SUI (n = 12) and UUI (n = 7)<br><br>Taught by: physiotherapist<br><br>Correct VPFMC confirmed? yes, by digital examination<br><br>Number VPFMC per set: progressive up to 23 contractions<br><br>Number sets per day: 5<br><br>Duration of hold: 10 s for sustained contraction, 2–3 s for short contractions<br><br>Duration of rest: 20 s for sustained contraction, 6 s for short contractions<br><br>Type(s) of contraction e.g. submaximal, maximal: sustained, maximal (short contractions) and submaximal (controlled)<br><br>Duration of programme: 8 weeks, 16 sessions (2 sessions per week)<br><br>Position: for all exercises position progressed from hook lying to sitting, standing squatting as able<br><br>Other important information on the intervention: none<br><br>**Group B:** control; SUI (n = 11) and UUI (n = 8)<br><br>Treatment postponed after follow-up<br><br>**Group C:** PFMT + electrical stimulation; SUI (n = 15) and UUI (n = 11) |
| Outcomes | **Primary endpoint:** post-intervention (8 weeks)<br><br>**Primary outcome:** QoL (YIPS)<br><br>**Secondary outcome:** pelvic muscle strength (perineometer), leaks and urination frequency (diary)<br><br>**Outcome measured during the intervention:** none reported<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none reported |
| Notes | **Dropouts and withdrawal:** total 5/38 (PFMT 1/19; control 4/19). No intention-to-treat analysis. |

**Firra 2013** *(Continued)*

**Funding:** "This study was funded in part by the Texas Physical Therapy Foundation."

**Conflicts of interest:** none declared

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Block randomisation not properly specified |
| Allocation concealment (selection bias) | Unclear risk | Not reported |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for the 'no treatment' arm |
| Blinding of outcome assessment (detection bias)<br>All outcomes | High risk | Not blinded |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Total of 5/38 (13%), unbalanced, reasons unknown<br>By group: PFMT 1/19 (5%); control 4/19 (21%)<br>No intention-to-treat analysis used |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | High risk | Not for 3-day frequency (control group urinated less frequently than the exercise group; P = 0.02) |

**Henalla 1989**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design<br>**Setting:** single centre, UK |
| Participants | 104 women with urodynamic SUI<br>**Method of diagnosis:** urodynamic<br>**Inclusion criteria:** urodynamically confirmed diagnosis of genuine SUI<br>**Exclusion criteria:** fistula, > 1 surgical procedure for incontinence, major degree of prolapse, absolute contraindication to oestrogens<br>**Mean age:** not reported |
| Interventions | **Group A (n = 26):** PFMT<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? yes, by vaginal palpation<br>Number VPFMC per set: 5 |

**Henalla 1989** *(Continued)*

| | |
|---|---|
| | Number sets per day: 1 per hour<br>Duration of hold: 5 s<br>Duration of rest: not reported<br>Type(s) of contraction e.g. submaximal, maximal: not reported<br>Duration of programme: 12 weeks<br>Body position: not reported<br>Other important information on the intervention:<br>Weekly clinic visit for 12 weeks<br>**Group B (n = 25):** control, no treatment<br>**Group C (n = 25):** electrical stimulation<br>**Group D (n = 24):** vaginal cream |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br>**Primary outcome:** not stated<br>**Other outcomes:** pad test cure (negative following positive result), pad test improvement (≥ 50% reduction in pad weight), cystometry<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** post-treatment evaluation at 12 weeks, with longer-term follow-up at 9 months (questionnaire) |
| Notes | **Dropouts and withdrawal:** none reported<br>**Funding:** not reported<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "were allocated at random" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate random allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding of the participants was not feasible for no treatment arm |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not clear if blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | No information about attrition |

**Henalla 1989** *(Continued)*

| | | |
|---|---|---|
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Quote: "the groups were comparable regarding age weight and parity." |

**Henalla 1990**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design<br>**Setting:** single centre, UK |
| Participants | 26 women with urodynamic SUI<br>**Inclusion criteria:** postmenopausal<br>**Exclusion criteria:** no further criteria stated<br>**Mean age (years):** 54 (range 49–64) |
| Interventions | **Group A (n = 8):** PFMT<br>No details given<br>**Group B (n = 7):** control, no treatment<br>**Group C (n = 11):** oestrogen therapy<br>**Group D (n = 22):** surgery |
| Outcomes | **Primary endpoint:** post-treatment evaluation (6 weeks)<br>**Primary outcome:** not stated<br>**Other outcomes:** pad test cure or improved (not defined), vaginal pH, vaginal cytology, anal EMG<br>**Any outcomes measured but not reported:** partial reporting of the outcomes presented in the method section due to brevity of abstract<br>**Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts:** not stated<br>**Funding:** not reported<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "randomised" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate random allocation concealment |
| Blinding of participants and personnel (performance bias) | Unclear risk | No details about control and treatment group blinding |

**Henalla 1990** *(Continued)*

| | | |
|---|---|---|
| All outcomes | | |
| Blinding of outcome assessment (detection bias) All outcomes | Unclear risk | Not clear if blinded outcome assessment |
| Incomplete outcome data (attrition bias) All outcomes | Unclear risk | No information about attrition |
| Selective reporting (reporting bias) | High risk | Partial reporting of the outcomes presented in the method section due to brevity of abstract |
| Baseline comparability | Unclear risk | Not clear if groups were comparable at baseline |

**Hofbauer 1990**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design |
| Participants | 43 women with urodynamic SUI<br>**Inclusion criteria:** not reported<br>**Exclusion criteria:** UUI<br>**Mean age (years):** 57.5 (SD 12)<br>Grade 3 incontinence: 4 PFMT; 2 control |
| Interventions | **Group A (n = 11):** PFMT<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? not reported<br>Number VPFMC per set: not reported<br>Number sets per day: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Type(s) of contraction e.g. submaximal, maximal: not reported<br>Duration of programme: 6 months<br>Body position: not reported<br>Other important information on the intervention:<br>Exercise programme including PFMT, abdominal and hip adductor exercise, twice per week for 20 min with therapist, and daily home programme<br>**Group B (n = 10):** control<br>Sham electrical stimulation<br>**Group C (n = 11):** PFMT + electrical stimulation<br>**Group D (n = 11):** electrical stimulation |

**Hofbauer 1990** *(Continued)*

| | |
|---|---|
| Outcomes | **Primary endpoint:** not clear when post-treatment evaluation performed<br>**Primary outcome:** not stated<br>**Other outcomes:** incontinence scale (not defined), leakage episodes (urinary diary), cystometry<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** further follow-up at 6 months |
| Notes | **Dropouts:** not stated<br>**Funding:** not reported<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "randomised" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if adequate random allocation concealment<br>Translated from German, "random" |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the treatment arm but there was sham electrical stimulation for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not clear if blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | No information about attrition |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Unclear risk | Not clear if groups were comparable at baseline |

**Kargar Jahromi 2013**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br>**Setting:** Jahandidegan Center, Kholdebarin Park in Shiraz, Iran |
| Participants | 48 women with SUI<br>**Method of diagnosis:** symptoms SUI (QUID questionnaire)<br>**Inclusion criteria:** women aged 60–74 years, QUID score for UI type (stress score ≥ 4), clinical symptoms of UI within the last 6 months, willing to participate in the study |

**Kargar Jahromi 2013** *(Continued)*

| | |
|---|---|
| | **Exclusion criteria:** absence in > 2 training sessions, central nervous system disease (e.g. multiple sclerosis, cerebrovascular accident or acute mental illness and dementia), recent urology surgery (for < 3 months), history of genitourinary malignancy, current urinary infection, hysterectomy and diabetic mellitus<br><br>**Mean age (years):** PFMT 67.15 (SD 8.36); control 68.05 (SD 9.1) |
| Interventions | **Group A (n = 24):** PFMT<br><br>Taught by: participants were thought to contract their PFM correctly, but without report of who did it<br><br>Correct VPFMC confirmed? no report on confirmation of correct contraction<br><br>Number VPFMC per set: 8–12 sustained + 4 fast contractions<br><br>Number sets per day: 3<br><br>Duration of hold: 6–8 s<br><br>Duration of rest: 6 s<br><br>Type(s) of contraction e.g. submaximal, maximal: maximal sustained and fast contractions<br><br>Duration of programme: 2 months. 8 group training classes (1 per week), 45 min<br><br>Position: lying, standing and sitting<br><br>Other important information on the intervention:<br><br>Education: the participants were taught about the anatomy of the pelvic floor and lower urinary tract, physiology, and continence mechanisms by the trained nurse + body awareness + breathing + relaxation<br><br>Home exercises: asked to be performed 3 times a day<br><br>**Group B (n = 24):** control, no details |
| Outcomes | **Primary endpoint:** post-treatment (2 months)<br><br>**Primary outcome:** ICIQ-UI Short Form, Rosenberg's self-esteem evaluation<br><br>**Other outcomes:** none<br><br>**Any outcomes measured but not reported:** none<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** total 2/50 (PFMT 1/25; control 1/25)<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Not reported |
| Allocation concealment (selection bias) | Unclear risk | Not reported |

**Kargar Jahromi 2013** *(Continued)*

| | | |
|---|---|---|
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group (no details about control group) |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not reported |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Total of 2/50 (4%)<br>By group: PFT 1/25 (4%) refused; control 1/25 (4%) migrated<br>No intention-to-treat analysis used |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | No difference between groups |

**Kim 2007**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, cross-over design<br>Stratification: level of physical fitness and leakage episode<br>Not clear if adequate random allocation concealment<br>Not clear if blinded outcome assessment<br>A priori power calculation<br>**Setting:** single urban centre, Japan |
| Participants | 70 women with SUI symptoms<br>**Method of diagnosis:** symptoms (ICIQ)<br>**Inclusion criteria:** urine leakage > once per month, UI associated with exertion<br>**Exclusion criteria:** UUI or MUI symptoms, no leakage or not enough<br>**Mean age (years):** PFMT 76.6 (SD 5.0); control 76.6 (SD 8.3)<br>Frequency score of urine leakage: PFMT 3.4 (SD 1.3); control 3.0 (1.3) |
| Interventions | **Group A (n = 35):** PFMT<br>Taught by: nurse<br>Correct VPFMC confirmed? not reported<br>Number VPFMC per set: 20<br>Number sets per day: 2 per week<br>Duration of hold: 3–8 s<br>Duration of rest: 10 s<br>Type(s) of contraction e.g. submaximal, maximal: sustained, fast |

**Kim 2007** *(Continued)*

| | |
|---|---|
| | Duration of programme: 12 weeks<br><br>Body position: sitting, lying and standing positions with the legs apart<br><br>Other important information on the intervention:<br><br>60-min exercise class twice per week for 12 weeks and 30-min home exercises twice per week<br><br>**Group B (n = 35):** control<br><br>Live normal life and refrain from exercises aiming to increase muscle strength, walking speed, to reduce BMI or to improve dietary habits for 12 weeks |
| Outcomes | **Primary endpoint:** post-treatment evaluation (3 months)<br><br>**Primary outcome:** ICIQ, frequency of UI leakage (worse to cured) at 3 and 12 months<br><br>**Other outcomes:** BMI, grip strength, walking speed, hip adductor strength<br><br>On a 6-point leakage scale of cure (0 = no urine leakage, 1 = < 1 per month, 2 = 1–2 per month, 3 = 1–2 per week, 4 = every 2 days and 5 = every day), the post-treatment score was significantly better for PFMT group than for the control group with a mean score post-treatment in the PFMT group of 1.5 (SD 1.8) and in the control group of 2.4 (SD 1.4) (MD –0.9, 95% CI –1.7 to –0.1)<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** longer-term follow-up at 12 months |
| Notes | **Dropouts:** total 5 (PFMT 2/35; control 3/35)<br><br>**Funding:** "This study was supported by a Research Grant of the Ministry of Health and Welfare of Japan and a Grant-in-Aid for Scientific Research B of the Japan Society for the Promotion of Science. Sponsor's Role: The sponsors had no role in the design of this study, development of the intervention program, subject recruitment, survey, data analysis, or preparation of the manuscript."<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "computer generated random number" |
| Allocation concealment (selection bias) | Unclear risk | Unclear: it was not clear how allocation concealment was conducted |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not clear if blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Quote: "5 participants ( 2 = PFMT and 3 = control group) where not able to complete study because of hospitalisation = 1, asthma =1, knee pain =1, or fracture = 2."<br><br>The reasons of attrition were not presented by group |

**Kim 2007** *(Continued)*

| | | |
|---|---|---|
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Groups comparable at baseline |

**Kim 2011a**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT<br><br>**Setting:** Basic Resident Register of 5935 women aged ≥ 70 years that resided in the Itabashi ward of Tokyo as of 1 April 2006 |
| Participants | 147 community-dwelling women aged ≥ 70 years with SUI, MUI or UUI<br><br>**Method of diagnosis:** symptoms<br><br>**Inclusion criteria:** urine leakage more than once per month; experiencing SUI, UUI and MUI according to symptoms; aged ≥ 70 years<br><br>**Exclusion criteria:** unclear type of UI, having urine leakage < 1 per month, impaired cognition (MMSE < 24), unstable cardiac conditions such as ventricular dysrhythmias, pulmonary oedema or other musculoskeletal conditions<br><br>**Mean age (years):** PFMT 76.7 (SD 3.6); control 75.8 (SD 3.6) |
| Interventions | **Group A (n = 37):** PFMT<br><br>Taught by: clinician giving the PFM and fitness protocol<br><br>Correct VPFMC confirmed? not reported<br><br>Number VPFMC per set: 10 fast and 10 sustained contractions<br><br>Number sets per day: 3<br><br>Duration of hold: 3 s for fast contractions, 6–8 s for sustained contractions<br><br>Duration of rest: 5 s for fast contractions and 10 s for sustained contractions<br><br>Type(s) of contraction e.g. submaximal, maximal: PFM contraction without excessively straining the abdomen, performed in lying, sitting, standing position with legs apart<br><br>Duration of programme: 60 min, twice per week for 12 weeks in groups<br><br>Number and type of contact with health professional(s): twice per week for 12 weeks<br><br>Other information: the participants were informed that straining the abdomen increases abdominal pressure and exerts pressure on the PFM. The participants were trained to exert force only on the PFM without excessively straining the abdomen<br><br>Warm-up and stretching 10–15 min including shoulder rotation, waist rotation and others, PFMT (as above) in addition to fitness: strength training of the thigh and abdominal muscles performed between PFMT, weightbearing exercises, ball exercises and others<br><br>Home exercises 2–3 sets of (PFM +13 exercises) ≥ 3 times per week (duration approximately 30 min)<br><br>**Group B (n = 36):** control<br><br>General education class once per month for 3 months where participants were educated on cognitive function, osteoporosis and oral hygiene<br><br>**Group C (n = 37):** PFMT + heat and steam |

**Kim 2011a** *(Continued)*

| | |
|---|---|
| | **Group D (n = 37):** heat and steam |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br>**Primary outcome:** subjective cure (interview) of urine loss episodes<br>Complete cessation of urine loss episode was defined as cured.<br>**Other outcomes:** functional fitness, change in frequency of urine loss episodes (5-point scale), ICIQ<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** not stated<br>**Funding:** "This research was supported in part by a Research Grant from the Ministry of Health and Welfare of Japan, a Grant-in-Aid for Scientific Research B from the Japanese Society for the Promotion of Science (19300236) and by the Sanitary Products Research Foundation of the Kao (Tokyo, Japan). Neither sponsors had any role in the design and conduct of the study; subject recruitment; collection, management, analysis and interpretation of data; or preparation of the manuscript."<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "computer-generated random numbers" |
| Allocation concealment (selection bias) | Low risk | Quote: "The investigators were blind to the allocation of interventions." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group but there was an inactive treatment |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not clear if blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Group A: PFMT 2/37 (5%)<br>Group B: control /36 (6%)<br>Reasons for not completing the study in all 4 cases not reported |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Groups comparable at baseline for anthropometric values, physical fitness, measures and interview survey |

**Kim 2011b**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT |

**Kim 2011b** *(Continued)*

| | |
|---|---|
| | **Setting:** urban community-based study |
| Participants | 127 community-dwelling women aged ≥ 70 years with SUI, MUI or UUI<br><br>**Method of diagnosis:** symptoms<br><br>**Inclusion criteria:** urine leakage more than once per week; experiencing SUI, UUI and MUI according to symptoms; aged ≥ 70 years; completing a 1-week urinary diary<br><br>**Exclusion criteria:** unclear type of UI; having urine leakage less than once per week; not completing the 1-week bladder diary; impaired cognition (MMSE < 24); unstable cardiac conditions such as ventricular dysrhythmias, pulmonary oedema or other musculoskeletal conditions<br><br>**Mean age (years):** PFMT 76.1 (SD 4.3); control 75.7 (SD 4.4) |
| Interventions | **Group A (n = 63):** PFMT<br><br>Taught by: clinician giving the PFM and fitness protocol<br><br>Correct VPFMC confirmed? not reported<br><br>Number VPFMC per set: 10 fast and 10 sustained contractions<br><br>Number sets per day: 3<br><br>Duration of hold: 3 s for fast contractions, 6–8 s for sustained contractions<br><br>Duration of rest: 5 s for fast contractions and 10 s for sustained contractions<br><br>Type(s) of contraction e.g. submaximal, maximal: PFM contraction without excessively straining the abdomen, performed in lying, sitting, standing position with legs apart<br><br>Duration of programme: 60 min, twice per week for 12 weeks in groups<br><br>Number and type of contact with health professional(s): twice per week for 12 weeks<br><br>Other important information on the intervention:<br><br>Warm-up and stretching 10–15 min, PFMT (as above) in addition to fitness: strength training of the thigh and abdominal muscles performed between PFMT, back, legs, trunk and use of an exercise ball<br><br>Home exercises: 2–3 sets of (PFM + 13 exercises) ≥ 3 times per week (duration approximately 30-min)<br><br>**Group B (n = 64):** control<br><br>General education class once per month for 3 months where participants were educated on cognitive function, osteoporosis and oral hygiene |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br><br>**Primary outcome:** cure rate of urine leakage episodes (bladder diary) at 3 and 7 months<br><br>**Other outcomes:** ICIQ frequency of UI leakage (scale 0–5) at 3 and 7 months, BMI, waist line, grip strength, walking speed, hip adductor strength<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** 7 months' follow-up |
| Notes | **Dropouts:** total 7/127 (PFMT 4/63; control 3/64)<br><br>**Funding:** "This research was supported in part by a Research Grant from the Ministry of Health and Welfare of Japan and a Grant-in-Aid for Scientific Research B from the Japan Society for the Promotion of Science (19300236) and was supported by the Sanitary Products Research Foundation of the KAO Corporation." |

**Kim 2011b** *(Continued)*

**Conflicts of interest:** none declared

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "computer-generated random number" |
| Allocation concealment (selection bias) | Low risk | Quote: "the randomisation procedure was blinded." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group but there was an inactive treatment |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Quote: "the investigators that evaluated the effects of the exercise treatment were blind to the allocation of interventions." |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Total: 7/127 (6%)<br>Group A: PFMT 4/63 (6%): hip fracture (n = 1), moving (n = 1), knee pain (n = 1), spouse care (n = 1)<br>Group B: control 3/64 (5%): death (n = 1) hospitalisation (n = 1), decreased motivation (n = 1) |
| Selective reporting (reporting bias) | Low risk | Reported on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Quote: "Most of the baseline characteristics were similar between the groups."<br>All those presented by the study were similar between groups |

**Lagro-Janssen 1991a**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br>**Setting:** 13 general practices, the Netherlands |
| Participants | 110 women with urodynamic SUI with or without DO<br>**Method of diagnosis:** urodynamic<br>**Inclusion criteria:** women aged 20–65 years reporting ≥ 2 leakage episodes per month<br>**Exclusion criteria:** previous incontinence surgery, neurological causes of incontinence, urinary tract infection, temporary cause of incontinence<br>**Mean age (years):** PFMT 46.1 (SD 10.1); control 44.6 (SD 8.2)<br>Symptoms for > 5 years: PFMT 55%; control 33%<br>Mean leakage episodes per 24 hours: PFMT 2.5 (SD 2.0); control 3.3 (SD 2.2)<br>**Diagnosis:** 66 (60%) urodynamic SUI, 20 (18%) MUI, 18 (16%) UUI, 6 (6%) other |

**Lagro-Janssen 1991a** *(Continued)*

| | |
|---|---|
| | Note: only data from women with urodynamic SUI are included in the review, because women with other diagnoses also had bladder training |
| Interventions | **Group A (n = 54, but 33 with urodynamic SUI only):** PFMT<br>Taught by: general practitioner<br>Correct VPFMC confirmed? yes, by vaginal palpation<br>Number VPFMC per set: 10<br>Number sets per day: 5–10<br>Duration of hold: 6 s<br>Duration of rest: not reported<br>Type(s) of contraction e.g. submaximal, maximal: not reported<br>Duration of programme: 12 weeks<br>Body position: not reported<br>Other important information on the intervention:<br>Advice about incontinence pads from practice assistant<br>Information on PFM function and how to contract by family doctor<br>**Group B (n = 56, but 33 with urodynamic SUI only):** control<br>Advice about incontinence pads only. Offered treatment after 12 weeks |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br>**Primary outcome:** not stated<br>**Other outcomes:** incontinence severity (12-point score), subjective assessment, health locus of control questionnaire, general health questionnaire, leakage episodes (7-day diary), self-reported treatment adherence, adverse events<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** longer-term follow-up at 6 months, 12 months and 5 years |
| Notes | **Dropouts:** total 4/110 (PFMT 1/54; control 3/56)<br>**Funding:** not reported<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | High risk | Consecutively (i.e. quasi-random because of alternation) |
| Allocation concealment (selection bias) | High risk | Quote: "the patient were assigned consecutively to the treatment or control groups which were stratified on the basis of the severity of their incontinence."<br>Inadequate allocation concealment |

**Lagro-Janssen 1991a** *(Continued)*

| | | |
|---|---|---|
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group but there was a non-active treatment |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Blinded outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | No dropout reported before 6 months (or end of study first phase, which was of interest for this review) |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Quote: "no significant difference were found" |

**Leong 2015**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br>No stratification<br>**Setting:** Six Elderly Health Centres (EHCs), Department of Health, Hong Kong |
| Participants | 55 women with UI<br>**Method of diagnosis:** clinical (medical officers in-charge)<br>**Inclusion criteria:** Chinese women aged > 65 years; clinical diagnosis of SUI, UUI, or MUI of a mild-to-moderate severity (based on the scoring system; Lagro-Janssen 1991b) made by the Elderly Health Centre medical officers in-charge<br>**Exclusion criteria:** active urinary tract infection, people taking diuretic medication, presence of bladder pathology, dysfunction due to genitourinary fistula, tumour, pelvic irradiation, neurological conditions or other chronic conditions (e.g. diabetes mellitus, Parkinson's disease), previous anti-incontinence surgery, significant cognitive impairment assessed by the Cantonese version of MMSE score, obesity (BMI > 30 kg/m$^2$), use of concomitant treatments during the trial<br>**Mean age (years):** PFMT 73.0 (SD 4.0); control 75.4 (SD 5.0) |
| Interventions | **Group A (n = 27): PFMT**<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? yes by digital palpation<br>Number VPFMC per set: progressive PFM exercise programme with aim to strengthen and focus on type I and type II fibres<br>Number sets per day: 3<br>Duration of hold: 5–10 s<br>Duration of rest: 10 s<br>Type(s) of contraction, e.g. submaximal, maximal: slow submaximal (5–30 times); maximal – fast (5–10 times); neuromuscular re-education (the 'knack') |

**Leong 2015** *(Continued)*

| | |
|---|---|
| | Duration of programme: 12 weeks/30 min<br><br>8 physiotherapy sessions, 30 min<br><br>Once per week for the first 4 weeks<br><br>Once every 2 weeks for the next 5–12 weeks<br><br>Position: progressive, from lying to standing<br><br>Other important information on the intervention:<br><br>Bladder training: strategies to increase the time interval between voids by a combination of progressive void schedules, urge suppression, distraction, self-monitoring and reinforcement<br><br>**Group B (n = 28): control**<br><br>Advice and educational pamphlet with information about management of UI at baseline (no active treatment and no contact with the therapist) |
| Outcomes | **Primary endpoint:** postintervention (12 weeks)<br><br>**Primary outcome:** UI episodes in the previous 7 days<br><br>**Secondary outcome:** QoL-IIQ-7, perception of improvement (VAS), adverse events, satisfaction to treatment (VAS): only intervention group<br><br>**Outcome measured during the intervention:** UI episodes; weekly until the end of the programme<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** 0<br><br>**Funding:** "This study was jointly funded by the Hong Kong Polytechnic University and the Department of Health)."<br><br>**Conflicts of interest:** not reported |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "Randomisation was performed prior to the study by an off-site investigator using a computerised randomisation programme." |
| Allocation concealment (selection bias) | Low risk | Quote: "allocation concealment by sequentially numbered, opaque, and sealed envelopes."<br><br>Quote: "After taking consent, grouping of the individual participants was revealed to the principal investigator by phone." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | High risk | Assessor same as therapist |

**Leong 2015** *(Continued)*

| | | |
|---|---|---|
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Reported, 0 losses |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section and trial registry available |
| Baseline comparability | High risk | Age: control group older (P = 0.06)<br>IIQ-7 score: control group have lower points (P = 0.06) (less impacted) |

**McLean 2013**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br>No stratification<br>**Setting:** Queen's University and Affiliated Hospitals Health Sciences |
| Participants | 35 women with SUI<br>**Method of diagnosis:** symptoms of SUI<br>**Inclusion criteria:** women on waiting list to see urogynaecologist for SUI complaints, aged ≥ 18 years, symptoms of SUI with or without UUI/nocturia/anterior compartment prolapse<br>**Exclusion criteria:** faecal incontinence, medications known to increase or alleviate incontinence, had known neurological impairments involving the central nervous system or the sacral nerves, known connective tissue disorders, detrusor instability on urodynamic, evidence of neurological defects, pelvic mass, prolapse greater than stage two (POP-Q)<br>**Mean age (years):** PFMT 49.5 (SD 8.2); control 54.0 (SD 8.4) |
| Interventions | **Group A (n = 18): PFMT**<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? yes, by digital palpation<br>Number VPFMC per set: 10 + specific exercises focused on holding or relaxing after the contraction<br>Number sets per day: 3<br>Duration of hold: ramp up over 2–4 s, hold 1–2 s<br>Duration of rest: not specified<br>Type(s) of contraction e.g. submaximal, maximal: maximum voluntary contraction and contract their PFMs before tasks that increase intra-abdominal pressure including coughing, laughing, sneezing and postural perturbations<br>Duration of programme: 12 weeks, 30 min per week<br>Position: not reported<br>Other important information on the intervention: none<br>**Group B (n = 17):** control, no intervention |
| Outcomes | **Primary endpoint:** postintervention (12 weeks) |

**McLean 2013** *(Continued)*

| | |
|---|---|
| | **Primary outcome:** 3-day bladder diary, 30-min pad test, IIQ-7, UDI-6<br><br>**Secondary outcome:** urethral morphometry (bladder neck position and mobility during coughing and Valsalva manoeuvre in supine and in standing + cross-sectional area)<br><br>**Outcome measured during the intervention:** modified Oxford scale during the 12 weeks of training (to provide feedback about progress) but not reported<br><br>**Any outcomes measured but not reported:** none except Oxford<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** total 5/40 (PFMT 2/20; control 3/20). No intention-to-treat analysis<br><br>**Funding:** none stated<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Custom automated algorithm |
| Allocation concealment (selection bias) | Unclear risk | Not reported |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not reported |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Total of 5/40 (13%), reasons unknown<br><br>By group: PFT 2/20 (10%); control 3/20 (15%)<br><br>No intention-to-treat analysis used<br><br>+ missing data from ultrasound images (5/20 for PFMT; 3/20 for control) |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | 2 groups comparable at baseline |

**Miller 1998**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design (after 1 month, controls cross over into treatment group)<br><br>**Setting:** single centre, USA |
| Participants | 27 women with of SUI<br><br>**Method of diagnosis:** symptoms and signs |

**Miller 1998** *(Continued)*

| | |
|---|---|
| | **Inclusion criteria:** community-dwelling women, mild-to-moderate SUI (≥ 1 and ≤ 5 leaks per day), aged ≥ 60 years, direct visualisation of urine loss on cough with 100 mL or more voided after stress test<br><br>**Exclusion criteria:** systemic neuromuscular disease, previous bladder surgery, active urinary tract infection, delayed leakage after cough, more than moderate leakage with cough, inability to do a VPFMC, prolapse below hymeneal ring<br><br>**Mean age (years):** 68.4 (SD 5.5)<br><br>**Mean number leakage episodes per day:** 1.4 (SD 1.4) |
| Interventions | **Group A (n = 13):** PFMT<br><br>Taught by: nurse<br><br>Correct VPFMC confirmed? yes, by vaginal palpation<br><br>Number VPFMC per set: not reported<br><br>Number sets per day: not reported<br><br>Duration of hold: not reported<br><br>Duration of rest: not reported<br><br>Type(s) of contraction e.g. submaximal, maximal: co-ordination and knack<br><br>Duration of programme: 1 week<br><br>Position: not reported<br><br>Other important information on the intervention:<br><br>Education on basic physiology and function of PFMs, digital palpation to teach VPFMC. Taught 'The Knack,' i.e. VPFMC prior to hard cough maintained throughout cough until abdominal wall relaxed. Practice at home for 1 week<br><br>**Group B (n = 14):** control<br><br>No treatment for 1 week, then cross-over to treatment group at 1 month |
| Outcomes | **Primary endpoint:** post-treatment (1 week)<br><br>**Primary outcome:** paper towel test<br><br>A paper towel test was reported as mean wet area and SD on either a moderate or a deep cough.<br><br>**Secondary outcome:** pelvic muscle strength (digital palpation)<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts and withdrawal:** none stated<br><br>**Funding:** "This work was supported by Public Health Service Grants P30 AG 08808,1101 DK 47516, and T32 AG 001140"<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|

**Miller 1998** *(Continued)*

| | | |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Quote: "randomly assigned in blocks of two" |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate allocation concealment. Quote: "randomly assigned in blocks of two" |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Blinding of outcome assessment |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Evaluation, only 1 week after and report on all participants |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Groups comparable at baseline |

**Pereira 2011**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design<br><br>**Setting:** single centre. Laboratory for assessment and intervention on Women's health, Federal university of Sao Carlos, Brazil |
| Participants | 49 women with SUI aged > 18 years<br><br>**Method of diagnosis:** SUI symptoms<br><br>**Inclusion criteria:** complain of urinary leakage on stress (2 standard questions about SUI and UUI used to determine participant eligibility: during the past month, have you involuntary got wet while performing some kind of physical exertion, coughing, lifting, sneezing or laughing? For urgency, the question was: during the past month, have you experienced such a strong urge to urinate that it was impossible to get to the toilet on time? Those answering yes to the stress question only and who had not undergone physical therapy for UI before were included)<br><br>**Exclusion criteria:** symptoms of UUI and MUI, latex allergies, vaginal or urinary infections, pelvic organ prolapse greater than grade II on Baden-Walker classification system, cognitive or neurological disorder, uncontrolled hypertension and inability to carry out the evaluation or treatment<br><br>**Mean age (years):** PFMT 60.20 (SD 8.16); individual PFMT 60.6 (SD 12.63); control 61.53 (SD 10.11) |
| Interventions | **Group A (n = 17):** group PFMT<br><br>Taught by: physical therapist<br><br>Correct VPFMC confirmed? yes with vaginal palpation<br><br>Number VPFMC per set: not clear, 100 in total on average in intervention sessions<br><br>Number sets per day: not mentioned |

**Pereira 2011** *(Continued)*

| | |
|---|---|
| | Duration of hold during intervention sessions: (mean time of the group was considered as the time of sustained contraction). The time of sustained contraction was increased by 1 s per week up to 10 s<br><br>Duration of rest during intervention sessions: double the duration of hold<br><br>Type(s) of contraction e.g. submaximal, maximal. Quote:"100 contractions were performed on average, composed of phasic contractions held for 3 s with 6 s rest and tonic contractions of 5–10 s followed by 10–20 s rest. To minimize the muscle fatigue, the resting time was rigidly observed in all sessions and the time of sustained contraction was slowly increased. PFMT was carried out in supine, sitting and standing positions. The degree of difficulty progressed according to the positions adopted, the number of repetitions, and the time of sustained contraction."<br><br>**Group B (n = 17):** individual PFMT:<br><br>Taught by: physical therapist<br><br>Correct VPFMC confirmed? yes with vaginal palpation<br><br>Number VPFMC per set: not clear, 100 in total on average in intervention sessions<br><br>Number sets per day: not mentioned<br><br>Duration of hold: 3–10 s during intervention sessions. The time of sustained contraction was increased by 1 s per week up to 10 s<br><br>Duration of rest: 6–20 s in intervention sessions<br><br>Type(s) of contraction, e.g. submaximal, maximal. Quote: "100 contractions were performed on average, composed of phasic contractions held for 3 s with 6 s rest and tonic contractions of 5–10 s followed by 10–20 s rest. To minimize the muscle fatigue, the resting time was rigidly observed in all sessions and the time of sustained contraction was slowly increased. PFMT was carried out in supine, sitting and standing positions. The degree of difficulty progressed according to the positions adopted, the number of repetitions, and the time of sustained contraction."<br><br>Other important information on the group and individual interventions:<br><br>Duration of programme: 2 × 1 hour weekly sessions in clinic for 6 weeks<br><br>Number and type of contact with health professional(s): 12 group or individual sessions twice per week for 1 hour for a total of 6 weeks<br><br>Explanation about anatomy of the PFM and continence mechanism<br><br>**Group C (n = 15):** control<br><br>Did not receive any treatment during the corresponding treatment time |
| Outcomes | **Primary endpoint:** post-treatment (6 months)<br><br>**Primary outcome:** urinary loss (1-hour pad test)<br><br>**Secondary outcome:** KHQ, PFM pressure perineometry, PFM strength by digital palpation and subjective satisfaction with treatment (the only 2 response options available were 'satisfied' and 'dissatisfied:' 'satisfied' indicated that the participant did not want a different treatment, 'dissatisfied' indicated that the participant wanted a different treatment from the initial one)<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** none |
| Notes | **Dropouts and withdrawal:** total 2/123 (PFMT 1/62; control 1/61). Intention-to-treat analysis<br><br>**Funding:** "The authors would like to acknowledge the funding support from Brazilian National Research Council (CNPq)." |

**Pereira 2011** *(Continued)*

**Conflicts of interest:** none declared

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "participants blindly drew one of the 49 preprinted cards in opaque sealed envelopes from a box."<br>No mention of successively numbered |
| Allocation concealment (selection bias) | Unclear risk | Quote: "participants blindly drew one of the 49 preprinted cards in opaque sealed envelopes from a box." |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | High risk | Evaluator was not blinded |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Total: 4/34 (8%)<br>Group intervention = 2/17 (12%)*<br>Individual intervention = 2/17 (12%)*<br>Control intervention = 0/15 0%<br>*Reasons: health problem or family (information not given per treatment group) |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section and trial registry available |
| Baseline comparability | Low risk | Group similar at baseline for demographics and clinic characteristics |

**Sar 2009**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br>**Setting:** 2 centres, outpatient urology clinics attached to a country hospital and a university hospital in Izmir, Turkey |
| Participants | 41 women with UI<br>**Method of diagnosis:** signs (2 g of urine on a 1-hour pad test)<br>**Inclusion criteria:** women with stress or mixed signs on surgical waiting list between 2005 and 2007, MMSE score: ≥ 25<br>**Exclusion criteria:** UTI, previous surgery of UI, neurological disease, diabetes mellitus, comorbid conditions likely to interfere with treatment, UI medication, inability to understand Turkish language<br>**Mean age (years)**: PFMT 41.82 (SD 8.65); control 44.64 (SD 6.90) |
| Interventions | **Group A (n = 19):** PFMT |

**Sar 2009** *(Continued)*

| | |
|---|---|
| | Taught by: nurse<br><br>Correct VPFMC confirmed? yes using vaginal palpation<br><br>Number VPFMC per set: 30 contractions per set<br><br>Number sets per day: 3<br><br>Duration of hold: 1–10 s<br><br>Duration of rest: same as contraction time<br><br>Type(s) of contraction e.g. submaximal, maximal: quick flicks (1–2 s contractions), sustained progressive (5–10 s) contractions + knack<br><br>Duration of programme: 6 weeks<br><br>Position: supine, sitting and standing<br><br>Other important information on the intervention:<br><br>Taught about the anatomy of the pelvic floor, lower urinary tract anatomy and continence mechanism.<br><br>Information was summarised in an illustrated handbook<br><br>**Group B (n = 22):** control, not contacted |
| Outcomes | **Primary endpoint:** post-treatment (6 weeks)<br><br>**Primary outcome:** not stated<br><br>**Other outcomes:** I-QOL; PFM vaginal squeeze pressure (perineometer); 3-day bladder diary; urine loss (1-hour pad test)<br><br>**Any outcomes measured but not reported:** none<br><br>**Follow-up after primary endpoint:** none |
| Notes | All outcomes were reported as change scores and SD, which we could not use in our forest plot. All outcomes significantly favoured PFMT vs control (P < 0.01)<br><br>I-QOL: PFMT 23.19 (SD 11.43); control 5.74 (SD 6.26)<br><br>Bladder diary (change in leakage/3 days): PFMT –3.23 (SD 2.19); control 0.82 (SD 2.81)<br><br>1-hour pad test (change in grams from baseline): PFMT –5.11 (SD 7.29); control 8.88 (SD 12.52)<br><br>PFM strength: mean and maximum as pressure using intravaginal perineometry: mean: PFMT 9.47 (SD 6.53); control –2.23 (SD 4.43); maximum: PFMT 11.23 (SD 7.60); control –3.70 (SD 4.71)<br><br>**Dropouts and withdrawal:** total 5/41 (PFMT 2/19; control 5/22). No intention-to-treat analysis.<br><br>**Funding:** not reported<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Quote: "randomly assigned to an intervention or control group"<br><br>Quote: "stratified based on PFM strength, frequency of UI episodes and severity of UI on a 1h pad test" |

**Sar 2009** *(Continued)*

| | | |
|---|---|---|
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | High risk | Quote: "this trial was not blinded." |
| Incomplete outcome data (attrition bias)<br>All outcomes | High risk | Total 7/41 (17%)<br>Group A = 2 (11%) dropout: non-adherence to treatment regimen<br>Group B = 5 (23%) excluded: used other treatment during the trial |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | No significant differences at baseline for age, BMI, parity, cystocoele, rectocoele duration of symptoms, menopause status, PFM strength, episode of leakage, 1-hour pad tests, I-QOL scores |

**Solberg 2016**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design (pilot trial)<br>**Setting:** single centre, Oslo, Norway |
| Participants | 20 women with MUI<br>**Method of diagnosis:** symptoms of MUI<br>**Inclusion criteria:** women, aged ≥ 18 years, MUI, not pregnant or planning to become pregnant during the study<br>**Exclusion criteria:** have given birth within 12 months using medication for UI, had undergone surgery for UI<br>**Median age (years):** PFMT 63.5 (range 40–87); control 52.5 (range 29–76) |
| Interventions | **Group A (n = 10):** PFMT<br>Taught by: physiotherapist<br>Correct VPFMC confirmed? Yes, using vaginal examination<br>Number VPFMC per set: 12<br>Number sets per day: 3<br>Duration of hold: 6 s<br>Duration of rest: at least 1 breath out and 1 in<br>Type(s) of contraction e.g. submaximal, maximal: maximal |

**Solberg 2016** *(Continued)*

| | |
|---|---|
| | Duration of programme: 12 weeks (25 min of PFMT + 20 min of general exercises, 1 per week + 10 min home PFMT/daily)<br>Position: not reported<br>Other important information on the intervention:<br>20 min of general exercise<br>**Group B (n = 12):** control<br>No treatment waiting list<br>**Group C (n = 12):** acupuncture |
| Outcomes | **Primary endpoint:** post-treatment (12 weeks)<br>**Primary outcome:** ICIQ-UI Short Form<br>**Secondary outcome:** adverse events, adherence<br>**Outcomes measured during the intervention:** not reported<br>**Any outcomes measured but not reported:** no<br>**Follow-up after primary endpoint:** no |
| Notes | Primary outcome different at baseline<br>Small sample size (6 instead of the 43 needed according to the sample size calculation)<br>High attrition rate (PFMT: from 10 to 6; control: from 12 to 6)<br>**Dropouts and withdrawal:** total 10/22 (PFMT 4/10; control 6/12). No intention-to-treat analysis<br>**Funding:** "Funding The Norwegian Acupuncture Association funded open access fee."<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Computer generated by SurveyMonkey |
| Allocation concealment (selection bias) | Low risk | Central allocation web-base |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Not reported |
| Incomplete outcome data (attrition bias)<br>All outcomes | High risk | Total of 10/22 (45%)<br>By group: PFT 4/10 (40%); control 6/12 (50%)<br>Unbalanced + high attrition rate in both groups |

**Solberg 2016** *(Continued)*

| | | |
|---|---|---|
| | | No intention-to-treat analysis used |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | High risk | Groups were not comparable at baseline for primary outcome (ICIQ-UI Short Form) |

**Sran 2016**

| | |
|---|---|
| Methods | **Design:** 2-arm RCT, parallel design<br><br>**Setting:** single centre, BC Women's Health Center in Vancouver, Canada |
| Participants | 48 women with UI<br><br>**Method of diagnosis:** signs and symptoms of SUI, UUI, or MUI<br><br>**Inclusion criteria:** postmenopausal with osteoporosis or low bone density; defined by a T score of –2.0 or lower for the lumbar spine or hip, or a history of a non-traumatic hip, vertebral, wrist or rib fracture; aged≥ 55 years; symptoms of SUI, UUI, or MUI for at least the past 3 months and ≥ 2 UI episodes in 3 days (self-reported); able to communicate in English (both written and verbal); willing to give written consent to participate<br><br>**Exclusion criteria:** previous treatments or workshops on incontinence in the past 5 years; previous UI surgeries (except for women who had had anti-incontinence surgery at least 20 years previously); faecal incontinence; continuous urine leakage; a current urinary tract infection; perineal pain or genital prolapse likely to interfere with the PFM assessment and treatment; previous pelvic irradiation; hormone therapy, use of vaginal oestrogen, or an unstable hormone dose within the previous 6 months; use of concomitant treatments for UI during the trial period; severe mobility impairments requiring the use of mobility aids (that would make going to the toilet difficult); use of high-dose diuretics or medications to improve bladder control; history of radiation for pelvic organ cancers; MMSE score < 24; any other medical problem likely to interfere with treatment and evaluation (serious cardiovascular disease, ongoing cancer treatments, neurological conditions, psychiatric conditions); and individuals performing a Valsalva manoeuvre in lieu of PFM contraction<br><br>**Mean age (years):** PFMT 66.17 (SD 6.66); control 67.13 (SD 8.38) |
| Interventions | **Group A (n = 24):** PFMT<br><br>Taught by: physiotherapist<br><br>Correct VPFMC confirmed by digital palpation<br><br>Number VPFMC per set: progressive<br><br>Number sets per day: 1<br><br>Duration of hold: individualised according to Laycock's perfect scheme<br><br>Duration of rest: not reported<br><br>Type(s) of contraction e.g. submaximal, maximal: fast MVC, slow MVC, knack (static dynamic), PFM + TA lying and standing, urge suppression techniques, crown (PFMs contracted 100%, dropped to 50% to 70%, up to 100%), PFME during walking, lunging and squatting<br><br>Duration of programme: 12 weeks. 1 hour for the first week, 30 min for 2nd to 12th<br><br>Position: lying, standing, walking, lunging and squatting<br><br>Other important information on the intervention: |

**Sran 2016** *(Continued)*

| | |
|---|---|
| | EMG biofeedback was used to monitor<br><br>**Group B (n = 24):** control<br><br>3-hour session in group on diet and osteoporosis + 2- to 4-hour individual follow-up by health professional |
| Outcomes | **Primary endpoint:** postintervention (12 weeks)<br><br>Number of leakage episodes on the 7-day bladder diary, pad test and disease-specific QoL and self-efficacy questionnaires<br><br>**Primary outcome:** number of leakage episodes on the 7-day bladder diary<br><br>**Secondary outcome:** pad test, UDI, IIQ, self-efficacy questionnaire<br><br>**Outcomes measured during the intervention:** no<br><br>**Any outcomes measured but not reported:** no<br><br>**Follow-up after primary endpoint:** at 1 year<br><br>Number of leakage episodes on the 7-day bladder diary, pad test and disease-specific QoL and self-efficacy questionnaires |
| Notes | UDI at 3 months data are presented unavailable through authors as mean and SD, only available as median and quartile<br><br>**Dropouts and withdrawal:** total 5/48 (PFMT 2/24; control 3/24). Intention-to-treat analysis<br><br>**Funding:** "Funding/support: This study was funded by the BC Women's Health Research Institute, Doris Winterbottom Research Award."<br><br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Low risk | Computer-generated random-block assignment |
| Allocation concealment (selection bias) | Low risk | Opaque, sequentially numbered, sealed envelopes |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group although they received an inactive treatment |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Low risk | Outcome assessors, data analysts, and the data collection and analysis team were blinded to group allocations |
| Incomplete outcome data (attrition bias)<br>All outcomes | Low risk | Total of 5/48 (10%)<br><br>By group: PFT 2/24 (8%); control 3/24 (13%)<br><br>Reasons for missing outcome data unlikely to be related to true outcome<br><br>Intention-to-treat analysis used |

**Sran 2016** *(Continued)*

| | | |
|---|---|---|
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section and registry available |
| Baseline comparability | Low risk | No statistical difference between groups |

**Wells 1999**

| | |
|---|---|
| Methods | **Design:** 4-arm RCT, parallel design<br><br>**Setting:** single centre, USA |
| Participants | 286 community living women, with SUI or MUI<br><br>**Method of diagnosis:** symptoms<br><br>**Inclusion criteria:** aged > 21 years, self-described as having uncontrolled urine loss or excessive day toileting frequency (or both), independent in self-care, able to speak and hear a conversation in English adequately over the telephone, negative urinalysis, able to contract the PFM as demonstrated on physical examination, able to read, understand and agree to the diagnostic consent form<br><br>**Exclusion criteria:** diagnosis of degenerative neurological disorder, pregnancy, high risk of infection following urological instrumentation<br><br>**Mean age (years):** 56 (SD 12.76) |
| Interventions | **Group A (n = 71):** PFMT<br><br>Taught by: nurse<br><br>Correct VPFMC confirmed? yes, by physical examination<br><br>Number VPFMC per set: 80<br><br>Number sets per day: 1<br><br>Duration of hold: 10 s<br><br>Duration of rest: 10 s<br><br>Type(s) of contraction e.g. submaximal, maximal: sustained<br><br>Duration of programme: 5 months<br><br>Position: not reported<br><br>Other important information on the intervention:<br><br>Initial training and active PFM exercises then monthly visits for observation, coaching and encouragement<br><br>**Group B (n = 72):** control<br><br>Directed 1 week per month to keep a daily record of fluid intake, toileting and urine leakage and discern a pattern and make simple lifestyle alterations if possible. Received diary by mail monthly<br><br>**Group C (n = 72):** health promotion<br><br>**Group D (n = 71):** resistive PFM exercise |
| Outcomes | **Primary endpoint:** post-treatment (5 months)<br><br>**Primary outcome:** not stated |

**Wells 1999** *(Continued)*

| | |
|---|---|
| | **Other outcomes:** PFM strength, urethral pressure and wetting |
| | **Any outcomes measured but not reported:** no |
| | **Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts and withdrawal:** total 65/143 (PFMT 30/71; controls 35/72). No intention-to-treat |
| | **Funding:** not reported |
| | **Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Participants were randomly assigned |
| Allocation concealment (selection bias) | Unclear risk | Not clear if there was adequate allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group although they received an inactive intervention |
| Blinding of outcome assessment (detection bias)<br>All outcomes | Unclear risk | Outcome assessment not blinded |
| Incomplete outcome data (attrition bias)<br>All outcomes | Unclear risk | Not clear if there was incomplete outcome data. No intention-to-treat analysis |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Unclear risk | Not clear if the groups were comparable at baseline |

**Yoon 2003**

| | |
|---|---|
| Methods | **Design:** 3-arm RCT, parallel design |
| | **Setting:** single centre, Korea |
| Participants | 46 women with UI |
| | **Method of diagnosis:** sign (urine loss ≥ 1.0 g on a 30-min pad test) |
| | **Inclusion criteria:** urine loss > 1 g on 30-min pad test, ≥ 14 voids in 48 hours |
| | **Exclusion criteria:** women aged < 35 and > 55 years, urinary tract infection, previous surgery for UI, hormonal or other drug therapy for UI |
| | **Mean age:** not reported |
| | Mean voids per day: PFMT 15.1 (SD 1.6); control 16.3 (SD 1.8) |

**Yoon 2003** *(Continued)*

| | |
|---|---|
| Interventions | **Group A (n = 12):** PFMT<br>Primary outcome: not stated<br>Taught by: nurse<br>Correct VPFMC confirmed? Yes, by EMG<br>Number VPFMC per set: 30 contractions per set<br>Number sets per day: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Type(s) of contraction e.g. submaximal, maximal: strength, endurance<br>Duration of programme: 8 weeks<br>Position: not reported<br>Other important information on the intervention: 20-min weekly session of EMG biofeedback with nurse, 8 weeks<br>**Group B (n = 13):** control<br>No treatment or clinic contact<br>**Group C (n = 19):** bladder training |
| Outcomes | **Primary endpoint:** post-treatment (8 weeks)<br>**Primary outcome:** not stated<br>**Other outcomes:** UI score (severity based on leakage with 18 activities), leakage episodes and frequency (2-day diary), 30-min pad test, vaginal squeeze pressure<br>**Any outcomes measured but not reported:** none<br>**Follow-up after primary endpoint:** no longer-term follow-up |
| Notes | **Dropouts and withdrawal:** total 4/29 (PFMT 2/15; controls 2/14). No intention-to-treat analysis<br>**Funding:** not reported<br>**Conflicts of interest:** none declared |

***Risk of bias***

| Bias | Authors' judgement | Support for judgement |
|---|---|---|
| Random sequence generation (selection bias) | Unclear risk | Using random number |
| Allocation concealment (selection bias) | Unclear risk | Quote: "assigned randomly to the control and treatment groups by using random numbers." Not clear if there was adequate allocation concealment |
| Blinding of participants and personnel (performance bias)<br>All outcomes | Unclear risk | Blinding was not possible for the control group |

**Yoon 2003** *(Continued)*

| | | |
|---|---|---|
| Blinding of outcome assessment (detection bias) All outcomes | Low risk | Blinded outcome assessment |
| Incomplete outcome data (attrition bias) All outcomes | Low risk | 2 women from the PFM group and 2 women from control withdrew due to family problems |
| Selective reporting (reporting bias) | Low risk | Reporting on all outcomes that were presented in the method section |
| Baseline comparability | Low risk | Quote: "no baseline difference" |

BMI: body mass index; CI: confidence interval; DO: detrusor overactivity; EMG: electromyography; I-QOL: Incontinence of Quality of Life questionnaire; ICIQ-UI Short Form: Incontinence Modular Questionnaire Urinary Incontinence Short Form; ICIQ-LUTSqol: ICIQ Lower Urinary Tract Symptoms Quality of Life; IEF: incontinence episode frequency; IIQ: Incontinence Impact Questionnaire; KHQ: King's Health Questionnaire; MD: mean difference; MESA: Medical Epidemiological and Social aspects of Aging; min: minute; MMSE: Mini-Mental State Examination; MUI: mixed urinary incontinence; MVC: maximal voluntary contraction; n: number of participants; OR: odds ratio; PDS: pressure drop under stress; PFM: pelvic floor muscle; PFMT: pelvic floor muscle training; PGI-I: Patient Global Impression of Improvement; POP-Q: Pelvic Organ Prolapse Quantifications; QoL: quality of life; QUID: Questionnaire for Urinary Incontinence Diagnosis; RCT: randomised controlled trial; s: second; SD: standard deviation; SUI: stress urinary incontinence; TA: transabdominal; UCP: urethral closure pressure; UDI: Urinary Distress Inventory; UI: urinary incontinence; UUI: urgency urinary incontinence; VAS: visual analogue scale; VPFMC: voluntary pelvic floor muscle contraction; YIPS: York Incontinence Perceptions Scale.

## Characteristics of excluded studies *[ordered by study ID]*

| Study | Reason for exclusion |
|---|---|
| Abdulaziz 2012 | 2-arm RCT comparing biofeedback assisted PFMT to a control group. Considered to be a comparison of PFMT and biofeedback to control |
| Albers-Heitner 2008 | Qualitative study, not an RCT |
| Albers-Heitner 2013 | Secondary analysis of a 2-arm RCT assessing satisfaction among women on the involvement of nurse specialists for patients with UI. Considered to be a comparison of combined treatments (not PFMT alone) vs general practice |
| Alves 2014 | 2-arm RCT comparing PFMT to physical fitness. The population included women with different types of urinary symptoms (e.g. pain, dryness, nodules), not only UI |
| Bernier 2008a | Electrical stimulation, biofeedback and PFMT used in the treatment arm of the RCT |
| Bernier 2008b | Electrical stimulation, biofeedback and PFMT used in the treatment arm of the RCT |
| Beuttenmuller 2011 | 3-arm RCT comparing PFMT, PFMT and Estim, and control. No UI outcome |
| Burgio 2002 | 3-arm RCT comparing PFMT and biofeedback, PFMT, and self-help booklet (including advice on PFMT). Considered to be a comparison of different approaches to PFMT |
| Chang 2011 | 3-arm RCT comparing acupressure, sham acupressure and usual care. No PFMT group |
| Dugan 2013 | 2-arm RCT comparing a fitness programme (Total Control Platinum Program - including PFMT) to education only. The intervention was considered to be a combination of PFMT and full-body, chair-based programme focusing on training the transversus abdominis and multifidus, and not PFM alone |

| Study | Reason for exclusion |
|---|---|
| Felicissimo 2010 | 2-arm RCT comparing 2 PFMT interventions: intensive supervised and unsupervised PFMT |
| Ferreira 2011a | 2-arm RCT comparing 2 PFMT interventions: home-based and supervised PFMT |
| Ferreira 2011b | Intervention: PFM educational group intervention not PFMT |
| Ghaderi 2016 | The 'control' group received active treatment that included 'exercise;' not PFM exercise but exercises. It was not a 'no treatment' or 'sham treatment' or 'control' treatment |
| Ghoniem 2005 | PFMT vs sham PFMT comparison was considered to be confounded by the choice of sham PFMT |
| Ghroubi 2015 | 2-arm RCT comparing PFMT vs weight loss. No control group |
| Goode 2003 | 3-arm RCT comparing PFMT + electrical stimulation, PFMT, and self-help booklet (including advice on PFMT). Considered to be a comparison of different approaches to PFMT |
| Kang 2015 | 2-arm RCT. Participants with many of the conditions listed in the exclusions for the review (i.e. neurological conditions) |
| Kumari 2008 | 2-arm RCT comparing PFMT and bladder training to the absence of treatment. Considered to be a comparison of 2 combined treatments (not PFMT alone) |
| Lapointe 2014 | Secondary analysis of 2-arm RCTs investigating if octogenarians benefited from health education as younger adults. Focus on health promotion. Considered to be a comparison of health education among octogenarians and younger adults |
| McLean 2015 | Secondary analysis of 2-arm RCTs comparing surgery and PFMT to surgery and pamphlet with PFMT exercises. Considered to be a comparison of 2 combined treatments (not PFMT alone) |
| Miller 2009 | 2-arm RCT comparing Knack instruction provided by video to video on food pyramid instruction. Both 10 minutes long. Not considered a PFMT programme |
| NCT02001714 | 2-arm RCT comparing a 1 time group behavioural class (2 h) to no treatment. Not considered to be a PFMT programme |
| NTR1114 | 2-arm preventive and therapeutic RCT comparing PFMT to control in women in early stage of cervical cancer with and without pelvic floor symptoms. Data of those with UI not presented separately |
| Ramsay 1990 | PFMT vs sham PFMT comparison that was considered to be confounded by the choice of sham PFMT |
| RBR-3fgwc7 | 3-arm RCT comparing PFMT, electrical stimulation and control. The control group included daily practice of PFMT, at home. Considered to be a comparison of different approaches to PFMT |
| Rutledge 2012 | 2-arm RCT comparing PFMT/behavioural therapy to usual care. Considered to be a comparison of a combined PFMT with bladder training intervention to control, not just PFMT alone. |
| Sjostrom 2015 | 2-arm RCT comparing 2 interventions, 1 by web and 1 by pamphlets; no control group |
| Tajiri 2014 | 2-arm RCT, control vs PFMT. However, the intervention included PFMT and transverse abdominal muscle cocontraction exercises, and not PFMT alone |
| Talley 2017 | 2-arm RCT comparing no treatment to a combination of PFMT and walking and strength training classes in frail older women with UI. Considered to be a combination of treatments (not PFMT alone) |

| Study | Reason for exclusion |
|---|---|
| van Leeuwen 2004 | 4-arm RCT comparing duloxetine alone, duloxetine and imitation PFMT, PFMT and placebo, and PFMT alone. Imitation PFMT and PFMT was considered to be a comparison of different approaches to PFMT (no control group) |
| Yang 2012 | 2-arm RCT comparing PFMT and biofeedback and control in gynaecology cancer survivors not specific to UI; Quote: "women who scored above two on of at least one of the bowel, bladder or sexual function questionnaires were selected." |
| Yoon 1999 | 3-arm, probably quasi-randomised trial, comparing PFMT, electrical stimulation, and control (not defined), for women with urodynamic SUI. This abstract contains no data; P values only |

h: hour; PFM: pelvic floor muscle; PFMT: pelvic floor muscle training; RCT: randomised controlled trial; SUI: stress urinary incontinence; UI: urinary incontinence.

### Characteristics of studies awaiting assessment *[ordered by study ID]*

#### Bali 2016

| | |
|---|---|
| Methods | Quantitative experimental design was used to study the effectiveness of Kegel exercise on women with urinary incontinence<br><br>Samples were selected by using non-probability, purposive sampling technique and randomly divided into 2 groups, i.e. urogenital questionnaire and incontinence questionnaire we acquired before and after the intervention (4 weeks) |
| Participants | 70 women with urinary incontinence selected from Doiwala<br><br>Experimental group (n = 34)<br><br>Control group (n = 35) |
| Interventions | Intervention was given to the experimental group, i.e. Kegel exercise for 4 weeks (3 cycles (1 cycle consisted 5 times) 3 times a day) |
| Outcomes | After 4 weeks of Kegel exercise, urogenital distress score was reduced from 1.7 (SD 0.65) to 1.0 (SD 0.42) and incontinence impact score from 1.6 (SD 0.63) to 1.1 (SD 0.40) in the experimental group ($P < 0.05$). In the control group, there was no significant decline |
| Notes | Awaiting clarification regarding scoring system used for the UDI and IIQ questionnaires in the trial. Authors contacted on 23 February 2018 |

#### Zhang 2015

| | |
|---|---|
| Methods | This study aimed to observe the effect of pelvic floor rehabilitation on SUI in elderly women. The effect was assessed after treatment |
| Participants | 86 elderly women with SUI were randomly allocated into training group (n = 43) and control group (n = 43) |
| Interventions | Training group: attended a pelvic floor rehabilitation course<br><br>Control group: received routine education only |
| Outcomes | In the training group, the symptoms of all participants were improved. The symptom improvement rates in women with mild SUI was 95.65% and moderate SUI was 81.25%. Their pelvic floor muscle |

| Zhang 2015 *(Continued)* | |
|---|---|
| | strength was 3.29 (SD 0.98) after treatment, which was higher than that of pretreatment (2.56 (SD 0.42)). Mean 1HPTV was 5.93 g (SD 2.05) before treatment and 4.16 g (SD 1.84) after treatment (P < 0.05). But I-QOL showed no significant difference between the 2 groups after treatment (P > 0.05) |
| Notes | Information taken from an abstract containing very little information. Study appeared to meet the inclusion criteria but some uncertainty remained. Full paper not yet published. We were unable to find the authors' contact details to ask for more information regarding their publication |

1HPTV: not defined by the authors, probably refers to 1 hour pad test values in grams of urine; IIQ: Incontinence Impact Questionnaire; n: number of participants; SD: standard deviation; SUI: stress urinary incontinence; UDI: Urinary Distress Inventory.

**Characteristics of ongoing studies** *[ordered by study ID]*

| NCT03097549 | |
|---|---|
| Trial name or title | Mobile app-treatment of mixed and urgency urinary incontinence in women – a randomized controlled study |
| Methods | 2-arm RCT comparing a non face-to-face programme for PFMT (smartphone app) to an inactive control group (lifestyle information) |
| Participants | Women with UUI or MUI, established via telephone interview |
| Interventions | Information about PFMT and bladder training for UUI and MUI. Information about psychological factors and lifestyle factors that might interfere with incontinence. Exercises for the pelvic floor muscles, bladder training and behaviours of avoidance due to fear of leakage, a statistic function and possibility to set reminders. Individual advice regarding lifestyle factors and training based on findings from the bladder diary and answers in the questionnaires |
| Outcomes | **Primary outcome measures:** International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form<br><br>**Secondary outcome measures:** International Consultation on Incontinence Modular Questionnaire Lower Urinary Tract Symptoms Quality of Life; change from baseline Incontinence Episode Frequency; number of incontinence episodes per 48 hours transferred to per week; International Consultation on Incontinence Modular Questionnaire Overactive Bladder; Incontinence Catastrophizing Scale; change from baseline; usage of incontinence aids at 15 weeks; Patient Global Impression of Improvement; participant satisfaction and self-rated question about if the current treatment was sufficient |
| Starting date | 31 March 2017 |
| Contact information | eva.samuelsson@umu.se |
| Notes | Information taken from the trial registry (22 February 2018) |

MUI: mixed urinary incontinence; PFMT: pelvic floor muscle training; RCT: randomised controlled trial; UUI: urgency urinary incontinence.

## DATA AND ANALYSES

**Comparison 1. Pelvic floor muscle training (PFMT) versus no treatment, placebo or control**

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1 Participant-perceived cure after treatment | 6 | | Risk Ratio (M-H, Fixed, 95% CI) | Subtotals only |
| 1.1 Stress urinary incontinence | 4 | 165 | Risk Ratio (M-H, Fixed, 95% CI) | 8.38 [3.68, 19.07] |
| 1.2 Urgency urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 1.3 Mixed urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 1.4 Urinary incontinence (all types) | 3 | 290 | Risk Ratio (M-H, Fixed, 95% CI) | 5.34 [2.78, 10.26] |
| 2 Participant-perceived cure or improvement after treatment | 5 | | Risk Ratio (M-H, Fixed, 95% CI) | Subtotals only |
| 2.1 Stress urinary incontinence | 3 | 242 | Risk Ratio (M-H, Fixed, 95% CI) | 6.33 [3.88, 10.33] |
| 2.2 Urgency urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 2.3 Mixed urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 2.4 Urinary incontinence (all types) | 2 | 166 | Risk Ratio (M-H, Fixed, 95% CI) | 2.39 [1.64, 3.47] |
| 3 Urinary incontinence-specific symptom measures (King's Health Questionnaire/severity measure after treatment) | 3 | | Mean Difference (IV, Fixed, 95% CI) | Subtotals only |
| 3.1 Stress urinary incontinence | 3 | 145 | Mean Difference (IV, Fixed, 95% CI) | -13.14 [-21.10, -5.18] |
| 3.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 3.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 3.4 Urinary incontinence (all types) | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 4 Urinary incontinence-specific quality of life measures (King's Health Questionnaire/incontinence impact after treatment) | 3 | | Mean Difference (IV, Random, 95% CI) | Subtotals only |
| 4.1 Stress urinary incontinence | 3 | 145 | Mean Difference (IV, Random, 95% CI) | -13.44 [-32.24, 5.35] |
| 4.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 4.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 4.4 Urinary incontinence (all types) | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 5 Urinary incontinence-specific quality of life measures (King's Health Questionnaire/physical limitation) | 3 | | Mean Difference (IV, Fixed, 95% CI) | Subtotals only |
| 5.1 Stress urinary incontinence | 3 | 145 | Mean Difference (IV, Fixed, 95% CI) | -11.89 [-20.55, -3.23] |
| 5.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 5.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 5.4 Urinary incontinence (all types) | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 6 Quality of life measures – not condition specific (King's Health Questionnaire/general health score) | 3 | | Mean Difference (IV, Fixed, 95% CI) | Subtotals only |
| 6.1 Stress urinary incontinence | 3 | 145 | Mean Difference (IV, Fixed, 95% CI) | 1.81 [-3.40, 7.03] |
| 6.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 6.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 6.4 Urinary incontinence (all types) | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 7 Urinary incontinence-specific symptom measures (Incontinence Modular Questionnaire Urinary Incontinence short form) | 4 | | Mean Difference (IV, Fixed, 95% CI) | Subtotals only |
| 7.1 Stress urinary incontinence | 3 | 196 | Mean Difference (IV, Fixed, 95% CI) | -3.45 [-4.39, -2.52] |
| 7.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 7.3 Mixed urinary incontinence | 1 | 12 | Mean Difference (IV, Fixed, 95% CI) | -3.97 [-7.85, -0.09] |
| 7.4 Urinary incontinence (all types) | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 8 Urinary incontinence-specific quality of life measures (Incontinence Modular Questionnaire Lower Urinary Tract Symptoms Quality of Life) | 1 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |
| 8.1 Stress urinary incontinence | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 8.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 8.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 8.4 Urinary incontinence (all types) | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 9 Urinary incontinence-specific symptom measures (Urinary Distress Inventory short form) | 2 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |
| 9.1 Stress urinary incontinence | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 9.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 9.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 9.4 Urinary incontinence (all types) | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 10 Urinary incontinence-specific quality of life measures (Incontinence Impact Questionnaire short form) | 3 | | Mean Difference (IV, Random, 95% CI) | Subtotals only |
| 10.1 Stress urinary incontinence | 1 | 35 | Mean Difference (IV, Random, 95% CI) | -19.7 [-30.63, -8.77] |
| 10.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 10.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 10.4 Urinary Incontinence (all types) | 2 | 176 | Mean Difference (IV, Random, 95% CI) | -7.54 [-14.70, -0.39] |
| 11 Urinary incontinence-specific quality of life measures (Incontinence Impact Questionnaire long form) | 1 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |
| 11.1 Stress urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 11.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 11.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 11.4 Urinary incontinence (all types) | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 12 Urinary incontinence-specific quality of life measures (Incontinence of Quality of Life questionnaire) | 2 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |
| 12.1 Stress urinary incontinence | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 12.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 12.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 12.4 Urinary incontinence (all types) | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 13 Participant-perceived cure at up to 1 year | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | Totals not selected |
| 13.1 Stress urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 13.2 Urgency urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 13.3 Mixed urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 13.4 Urinary incontinence (all types) | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 14 Participant-perceived cure or improvement at up to 1 year | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | Totals not selected |
| 14.1 Stress urinary incontinence | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 14.2 Urgency urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 14.3 Mixed urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 14.4 Urinary incontinence (all types) | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 15 Urinary incontinence-specific symptom measures at 1 year (Urinary Distress Inventory long form) | 1 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |
| 15.1 Stress urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 15.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 15.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 15.4 Urinary incontinence (all types) | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 16 Urinary incontinence-specific quality of life measures at 1 year (Incontinence Impact Questionnaire long form) | 1 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 16.1 Stress urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 16.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 16.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 16.4 Urinary incontinence (all types) | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 17 Participant-perceived satisfaction | 3 | | Risk Ratio (M-H, Fixed, 95% CI) | Subtotals only |
| 17.1 Stress urinary incontinence | 2 | 105 | Risk Ratio (M-H, Fixed, 95% CI) | 5.32 [2.63, 10.74] |
| 17.2 Urgency urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 17.3 Mixed urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 17.4 Urinary incontinence (all types) | 1 | 108 | Risk Ratio (M-H, Fixed, 95% CI) | 2.77 [1.74, 4.41] |
| 18 Perception of improvement (visual analogue scale) | 1 | | Mean Difference (IV, Fixed, 95% CI) | Totals not selected |
| 18.1 Stress urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 18.2 Urgency urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 18.3 Mixed urinary incontinence | 0 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 18.4 Urinary incontinence (all types) | 1 | | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 19 Number of women needing further treatment | 2 | | Risk Ratio (M-H, Fixed, 95% CI) | Totals not selected |
| 19.1 Stress urinary incontinence | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 19.2 Urgency urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 19.3 Mixed urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 19.4 Urinary incontinence (all types) | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 20 Number of leakage episodes in 24 hours | 11 | | Mean Difference (IV, Random, 95% CI) | Subtotals only |
| 20.1 Stress urinary incontinence | 7 | 432 | Mean Difference (IV, Random, 95% CI) | -1.23 [-1.78, -0.68] |

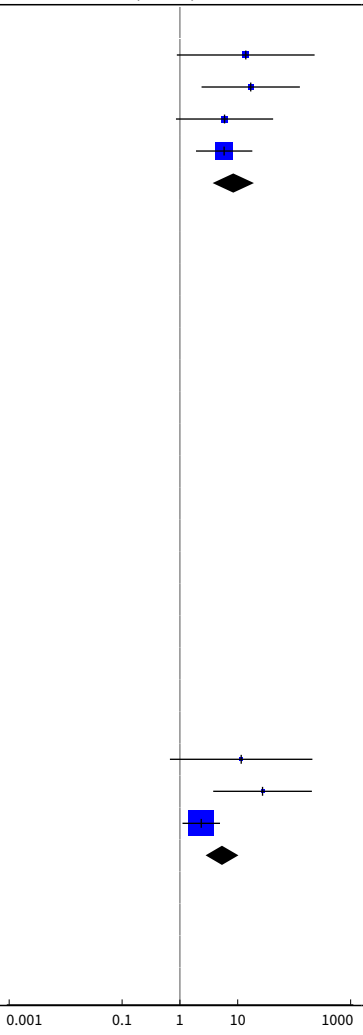
| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 20.2 Urge urinary incontinence | 1 | 12 | Mean Difference (IV, Random, 95% CI) | -1.83 [-2.65, -1.01] |
| 20.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 20.4 Urinary incontinence (all types) | 4 | 349 | Mean Difference (IV, Random, 95% CI) | -1.00 [-1.37, -0.64] |
| 21 Number of micturitions during the day (frequency) | 4 | | Mean Difference (IV, Random, 95% CI) | Subtotals only |
| 21.1 Stress urinary incontinence | 1 | 21 | Mean Difference (IV, Random, 95% CI) | -0.60 [-3.09, 1.89] |
| 21.2 Urgency urinary incontinence | 1 | 12 | Mean Difference (IV, Random, 95% CI) | -0.24 [-3.43, 2.95] |
| 21.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 21.4 Urinary incontinence (all types) | 3 | 187 | Mean Difference (IV, Random, 95% CI) | -2.32 [-3.21, -1.43] |
| 22 Number of micturitions during the night (nocturia) | 3 | | Mean Difference (IV, Fixed, 95% CI) | Subtotals only |
| 22.1 Stress urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 22.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 22.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 22.4 Urinary incontinence (all types) | 3 | 187 | Mean Difference (IV, Fixed, 95% CI) | -0.03 [-0.46, 0.40] |
| 23 Short (up to 1 hour) pad test measured as grams of urine | 6 | | Mean Difference (IV, Random, 95% CI) | Subtotals only |
| 23.1 Stress urinary incontinence | 4 | 185 | Mean Difference (IV, Random, 95% CI) | -9.71 [-18.92, -0.50] |
| 23.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 23.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Random, 95% CI) | 0.0 [0.0, 0.0] |
| 23.4 Urinary incontinence (all types) | 2 | 146 | Mean Difference (IV, Random, 95% CI) | -3.72 [-5.46, -1.98] |
| 24 Long (24 hours) pad test measured as grams of urine | 3 | | Mean Difference (IV, Fixed, 95% CI) | Subtotals only |
| 24.1 Stress urinary incontinence | 1 | 55 | Mean Difference (IV, Fixed, 95% CI) | -27.5 [-61.24, 6.24] |
| 24.2 Urgency urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 24.3 Mixed urinary incontinence | 0 | 0 | Mean Difference (IV, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 24.4 Urinary incontinence (all types) | 2 | 89 | Mean Difference (IV, Fixed, 95% CI) | -5.89 [-18.23, 6.44] |
| 25 Number cured on short pad test (objective) after treatment | 3 | | Risk Ratio (M-H, Fixed, 95% CI) | Subtotals only |
| 25.1 Stress urinary incontinence | 3 | 135 | Risk Ratio (M-H, Fixed, 95% CI) | 7.50 [2.89, 19.47] |
| 25.2 Urgency urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 25.3 Mixed urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 25.4 Urinary incontinence (all types) | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 26 Number cured or improved on short pad test (objective) after treatment | 3 | | Risk Ratio (M-H, Fixed, 95% CI) | Subtotals only |
| 26.1 Stress urinary incontinence | 3 | 96 | Risk Ratio (M-H, Fixed, 95% CI) | 8.22 [3.17, 21.28] |
| 26.2 Urgency urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 26.3 Mixed urinary incontinence | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 26.4 Urinary incontinence (all types) | 0 | 0 | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 27 Number of women with sex life spoilt by urinary incontinence | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | Totals not selected |
| 27.1 Stress urinary incontinence | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 27.2 Urgency urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 27.3 Mixed urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 27.4 Urinary incontinence (all types) | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 28 Number of women with urinary incontinence during intercourse | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | Totals not selected |
| 28.1 Stress urinary incontinence | 1 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 28.2 Urgency urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |

| Outcome or subgroup title | No. of studies | No. of partici-pants | Statistical method | Effect size |
|---|---|---|---|---|
| 28.3 Mixed urinary incontinence | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |
| 28.4 Urinary incontinence (all types) | 0 | | Risk Ratio (M-H, Fixed, 95% CI) | 0.0 [0.0, 0.0] |

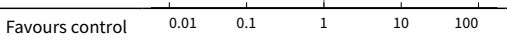
**Analysis 1.1. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 1 Participant-perceived cure after treatment.**

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| **1.1.1 Stress urinary incontinence** | | | | | |
| Hofbauer 1990 | 7/11 | 0/10 | | 9.38% | 13.75[0.88,213.65] |
| Bø 1999 | 14/25 | 1/30 | | 16.35% | 16.8[2.37,119.04] |
| Kim 2011a | 7/13 | 1/11 | | 19.48% | 5.92[0.86,41.03] |
| Kim 2007 | 18/33 | 3/32 | | 54.78% | 5.82[1.9,17.86] |
| **Subtotal (95% CI)** | **82** | **83** | | **100%** | **8.38[3.68,19.07]** |
| Total events: 46 (PFMT), 5 (Control) | | | | | |
| Heterogeneity: Tau²=0; Chi²=1.14, df=3(P=0.77); I²=0% | | | | | |
| Test for overall effect: Z=5.06(P<0.0001) | | | | | |
| **1.1.2 Urgency urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.1.3 Mixed urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.1.4 Urinary incontinence (all types)** | | | | | |
| Kim 2011a | 5/22 | 0/23 | | 5.13% | 11.48[0.67,196.07] |
| Kim 2011b | 26/59 | 1/61 | | 10.31% | 26.88[3.77,191.79] |
| Burgio 1998 | 19/63 | 8/62 | | 84.56% | 2.34[1.11,4.94] |
| **Subtotal (95% CI)** | **144** | **146** | | **100%** | **5.34[2.78,10.26]** |
| Total events: 50 (PFMT), 9 (Control) | | | | | |
| Heterogeneity: Tau²=0; Chi²=7.56, df=2(P=0.02); I²=73.55% | | | | | |
| Test for overall effect: Z=5.02(P<0.0001) | | | | | |
| Test for subgroup differences: Chi²=0.71, df=1 (P=0.4), I²=0% | | | | | |



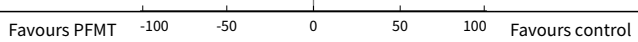
## Analysis 1.2. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 2 Participant-perceived cure or improvement after treatment.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| **1.2.1 Stress urinary incontinence** | | | | | |
| Asklund 2017 | 56/61 | 12/60 | | 86.37% | 4.59[2.75,7.66] |
| Bø 1999 | 12/25 | 1/30 | | 6.49% | 14.4[2.01,103.23] |
| Lagro-Janssen 1991a | 20/33 | 1/33 | | 7.14% | 20[2.85,140.51] |
| **Subtotal (95% CI)** | **119** | **123** | | **100%** | **6.33[3.88,10.33]** |
| Total events: 88 (PFMT), 14 (Control) | | | | | |
| Heterogeneity: Tau²=0; Chi²=3.52, df=2(P=0.17); I²=43.18% | | | | | |
| Test for overall effect: Z=7.38(P<0.0001) | | | | | |
| **1.2.2 Urgency urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.2.3 Mixed urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.2.4 Urinary incontinence (all types)** | | | | | |
| Burgio 1998 | 46/63 | 20/62 | | 85.69% | 2.26[1.53,3.35] |
| Diokno 2010 | 12/23 | 3/18 | | 14.31% | 3.13[1.04,9.45] |
| **Subtotal (95% CI)** | **86** | **80** | | **100%** | **2.39[1.64,3.47]** |
| Total events: 58 (PFMT), 23 (Control) | | | | | |
| Heterogeneity: Tau²=0; Chi²=0.3, df=1(P=0.58); I²=0% | | | | | |
| Test for overall effect: Z=4.57(P<0.0001) | | | | | |
| Test for subgroup differences: Chi²=9.62, df=1 (P=0), I²=89.61% | | | | | |

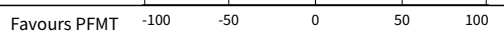



## Analysis 1.3. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 3 Urinary incontinence-specific symptom measures (King's Health Questionnaire/severity measure after treatment).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.3.1 Stress urinary incontinence** | | | | | | | |
| Beuttenmuller 2010 | 25 | 28.5 (24.1) | 25 | 36 (23.6) | | 36.33% | -7.45[-20.65,5.75] |
| Carneiro 2010 | 25 | 26.7 (26.7) | 25 | 34.6 (23.9) | | 32.02% | -7.95[-22.01,6.11] |
| Pereira 2011 | 30 | 20.9 (22.3) | 15 | 45.8 (23.1) | | 31.65% | -24.92[-39.06,-10.78] |
| **Subtotal ***** | **80** | | **65** | | | **100%** | **-13.14[-21.1,-5.18]** |
| Heterogeneity: Tau²=0; Chi²=3.9, df=2(P=0.14); I²=48.75% | | | | | | | |
| Test for overall effect: Z=3.24(P=0) | | | | | | | |
| **1.3.2 Urgency urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |

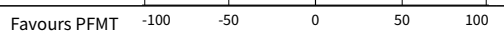


| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| Test for overall effect: Not applicable | | | | | | | |
| **1.3.3 Mixed urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.3.4 Urinary incontinence (all types)** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| Test for subgroup differences: Not applicable | | | | | | | |



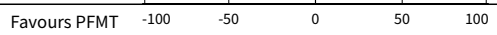

## Analysis 1.4. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 4 Urinary incontinence-specific quality of life measures (King's Health Questionnaire/incontinence impact after treatment).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Random, 95% CI | Weight | Mean Difference Random, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.4.1 Stress urinary incontinence** | | | | | | | |
| Beuttenmuller 2010 | 25 | 49 (26.8) | 25 | 53.8 (24.1) | | 35.13% | -4.8[-18.93,9.33] |
| Carneiro 2010 | 25 | 52.7 (28.7) | 25 | 55.4 (28.1) | | 33.5% | -2.75[-18.51,13.01] |
| Pereira 2011 | 30 | 23.3 (27.7) | 15 | 57.8 (29.5) | | 31.37% | -34.54[-52.44,-16.64] |
| **Subtotal ***** | **80** | | **65** | | | **100%** | **-13.44[-32.24,5.35]** |
| Heterogeneity: $Tau^2$=209.74; $Chi^2$=8.41, df=2(P=0.01); $I^2$=76.22% | | | | | | | |
| Test for overall effect: Z=1.4(P=0.16) | | | | | | | |
| **1.4.2 Urgency urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.4.3 Mixed urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.4.4 Urinary incontinence (all types)** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| Test for subgroup differences: Not applicable | | | | | | | |



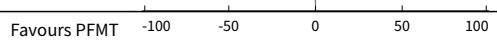
## Analysis 1.5. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 5 Urinary incontinence-specific quality of life measures (King's Health Questionnaire/physical limitation).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.5.1 Stress urinary incontinence** | | | | | | | |
| Beuttenmuller 2010 | 25 | 21.3 (19.5) | 25 | 31 (25.9) | | 46.51% | -9.72[-22.42,2.98] |
| Carneiro 2010 | 25 | 21.1 (26.6) | 25 | 29.2 (28) | | 32.66% | -8.05[-23.21,7.11] |
| Pereira 2011 | 30 | 7.2 (10) | 15 | 30 (36.8) | | 20.84% | -22.75[-41.73,-3.77] |
| **Subtotal ***** | **80** | | **65** | | | **100%** | **-11.89[-20.55,-3.23]** |
| Heterogeneity: Tau²=0; Chi²=1.62, df=2(P=0.45); I²=0% | | | | | | | |
| Test for overall effect: Z=2.69(P=0.01) | | | | | | | |
| **1.5.2 Urgency urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.5.3 Mixed urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.5.4 Urinary incontinence (all types)** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| Test for subgroup differences: Not applicable | | | | | | | |




## Analysis 1.6. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 6 Quality of life measures – not condition specific (King's Health Questionnaire/general health score).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.6.1 Stress urinary incontinence** | | | | | | | |
| Beuttenmuller 2010 | 25 | 29.3 (15.6) | 25 | 28.4 (16) | | 35.45% | 0.92[-7.84,9.68] |
| Carneiro 2010 | 25 | 34 (14.2) | 25 | 28.8 (14.7) | | 42.38% | 5.25[-2.76,13.26] |
| Pereira 2011 | 30 | 30 (15.6) | 15 | 33.3 (18.9) | | 22.18% | -3.33[-14.4,7.74] |
| **Subtotal ***** | **80** | | **65** | | | **100%** | **1.81[-3.4,7.03]** |
| Heterogeneity: Tau²=0; Chi²=1.58, df=2(P=0.45); I²=0% | | | | | | | |
| Test for overall effect: Z=0.68(P=0.5) | | | | | | | |
| **1.6.2 Urgency urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.6.3 Mixed urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |



| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.6.4 Urinary incontinence (all types)** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| Test for subgroup differences: Not applicable | | | | | | | |

Favours PFMT -100 -50 0 50 100 Favours control

## Analysis 1.7. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 7 Urinary incontinence-specific symptom measures (Incontinence Modular Questionnaire Urinary Incontinence short form).

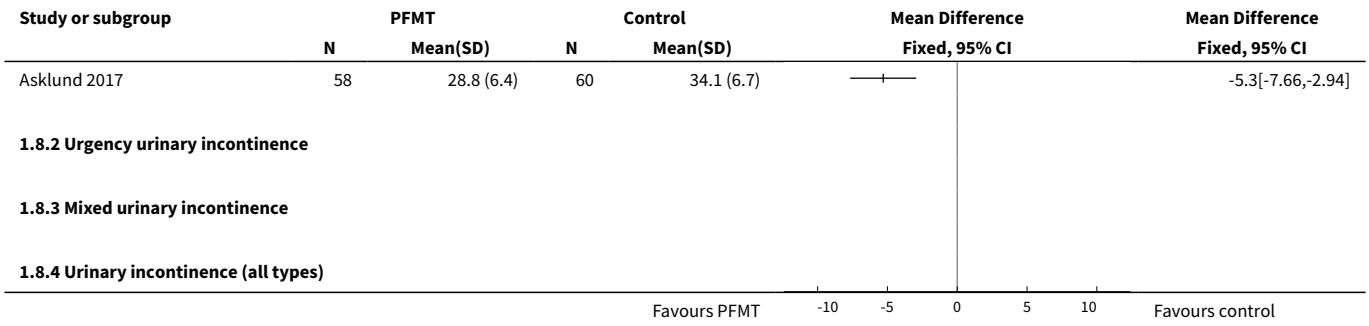
| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Weight | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.7.1 Stress urinary incontinence** | | | | | | | |
| Asklund 2017 | 59 | 7 (3.5) | 60 | 10.2 (3.2) | | 60.24% | -3.2[-4.41,-1.99] |
| Bertotto 2017 | 15 | 4.3 (3.2) | 14 | 10 (4.8) | | 9.79% | -5.7[-8.69,-2.71] |
| Kargar Jahromi 2013 | 24 | 9.1 (2.3) | 24 | 12.3 (3.6) | | 29.97% | -3.23[-4.94,-1.52] |
| **Subtotal ***** | **98** | | **98** | | | **100%** | **-3.45[-4.39,-2.52]** |
| Heterogeneity: $Tau^2$=0; $Chi^2$=2.4, df=2(P=0.3); $I^2$=16.77% | | | | | | | |
| Test for overall effect: Z=7.23(P<0.0001) | | | | | | | |
| **1.7.2 Urgency urinary incontinence** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| **1.7.3 Mixed urinary incontinence** | | | | | | | |
| Solberg 2016 | 6 | 5.8 (3.4) | 6 | 9.8 (3.4) | | 100% | -3.97[-7.85,-0.09] |
| **Subtotal ***** | **6** | | **6** | | | **100%** | **-3.97[-7.85,-0.09]** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Z=2(P=0.04) | | | | | | | |
| **1.7.4 Urinary incontinence (all types)** | | | | | | | |
| **Subtotal ***** | **0** | | **0** | | | | **Not estimable** |
| Heterogeneity: Not applicable | | | | | | | |
| Test for overall effect: Not applicable | | | | | | | |
| Test for subgroup differences: $Chi^2$=0.06, df=1 (P=0.8), $I^2$=0% | | | | | | | |

Favours PFMT -20 -10 0 10 20 Favours control

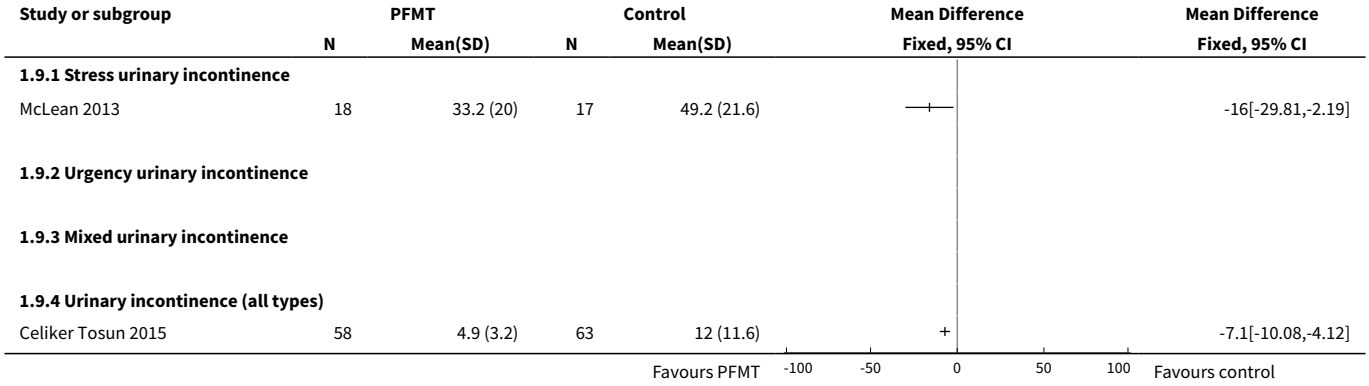
## Analysis 1.8. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 8 Urinary incontinence-specific quality of life measures (Incontinence Modular Questionnaire Lower Urinary Tract Symptoms Quality of Life).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|
| **1.8.1 Stress urinary incontinence** | | | | | | |

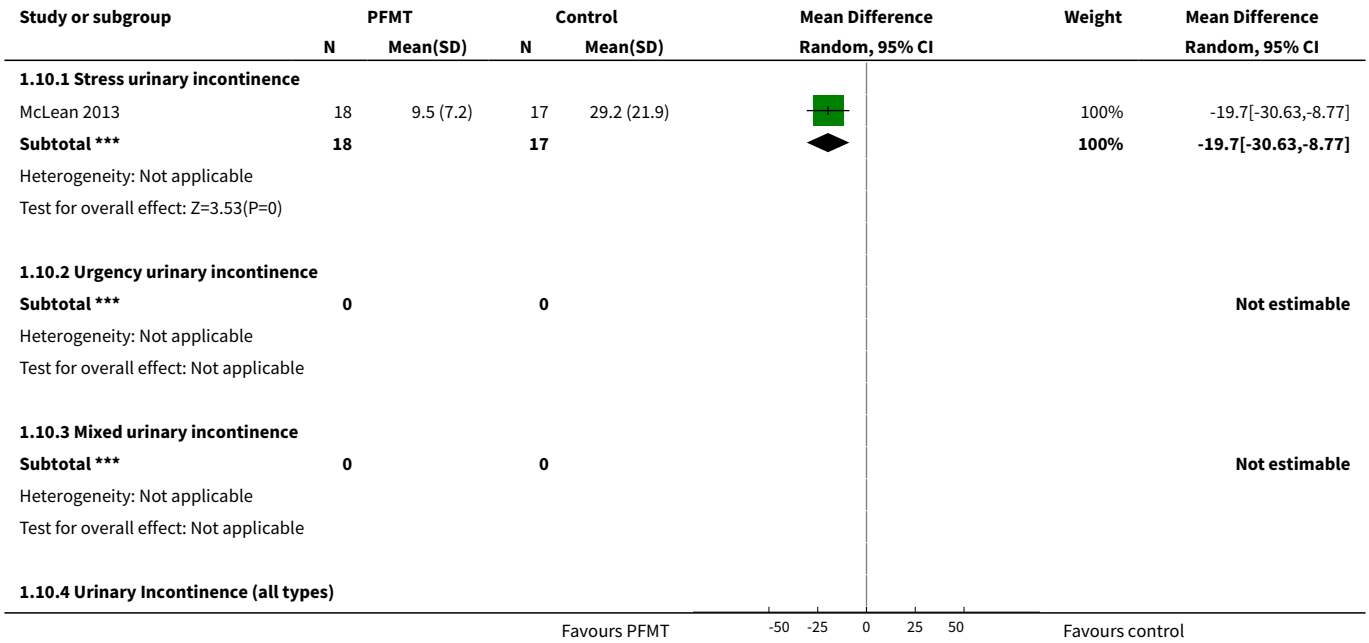
Favours PFMT -10 -5 0 5 10 Favours control




## Analysis 1.9. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 9 Urinary incontinence-specific symptom measures (Urinary Distress Inventory short form).




## Analysis 1.10. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 10 Urinary incontinence-specific quality of life measures (Incontinence Impact Questionnaire short form).



| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Random, 95% CI | Weight | Mean Difference Random, 95% CI |
|---|---|---|---|---|---|---|---|
| Celiker Tosun 2015 | 58 | 5.3 (4.3) | 63 | 16.5 (2.6) | | 49.91% | -11.2[-12.48,-9.92] |
| Leong 2015 | 27 | 1.1 (1.2) | 28 | 5 (2.8) | | 50.09% | -3.9[-5.03,-2.77] |
| **Subtotal ***** | **85** | | **91** | | | **100%** | **-7.54[-14.7,-0.39]** |
| Heterogeneity: Tau²=26.27; Chi²=70.17, df=1(P<0.0001); I²=98.57% | | | | | | | |
| Test for overall effect: Z=2.07(P=0.04) | | | | | | | |

Favours PFMT -50 -25 0 25 50 Favours control

## Analysis 1.11. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 11 Urinary incontinence-specific quality of life measures (Incontinence Impact Questionnaire long form).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|
| **1.11.1 Stress urinary incontinence** | | | | | | |
| **1.11.2 Urgency urinary incontinence** | | | | | | |
| **1.11.3 Mixed urinary incontinence** | | | | | | |
| **1.11.4 Urinary incontinence (all types)** | | | | | | |
| Sran 2016 | 24 | 26.5 (37.3) | 24 | 79.2 (99) | | -52.67[-95,-10.34] |

Favours PFM -100 -50 0 50 100 Favours control

## Analysis 1.12. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 12 Urinary incontinence-specific quality of life measures (Incontinence of Quality of Life questionnaire).

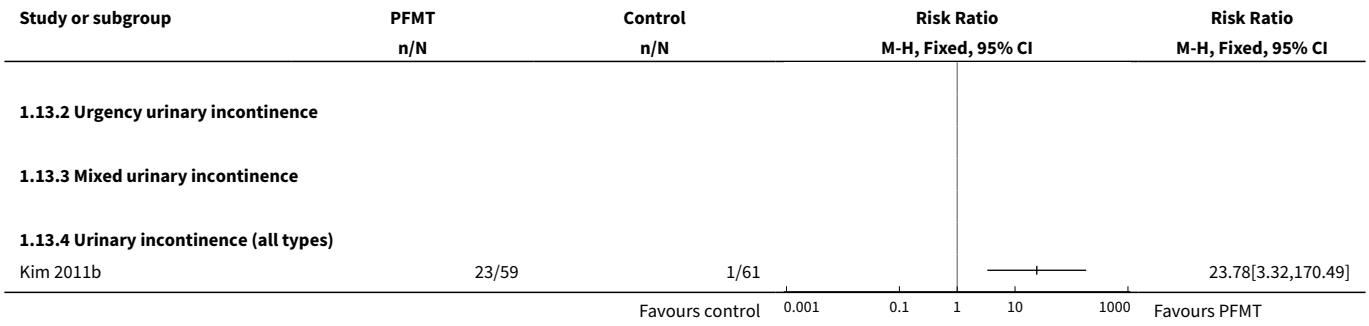
| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|
| **1.12.1 Stress urinary incontinence** | | | | | | |
| Castro 2008 | 26 | -82.2 (17.6) | 24 | -57.6 (28.2) | | -24.6[-37.75,-11.45] |
| **1.12.2 Urgency urinary incontinence** | | | | | | |
| **1.12.3 Mixed urinary incontinence** | | | | | | |
| **1.12.4 Urinary incontinence (all types)** | | | | | | |
| Sar 2009 | 17 | -23.2 (11.4) | 17 | 5.7 (6.3) | | -28.93[-35.12,-22.74] |

Favours PFMT -50 -25 0 25 50 Favours control

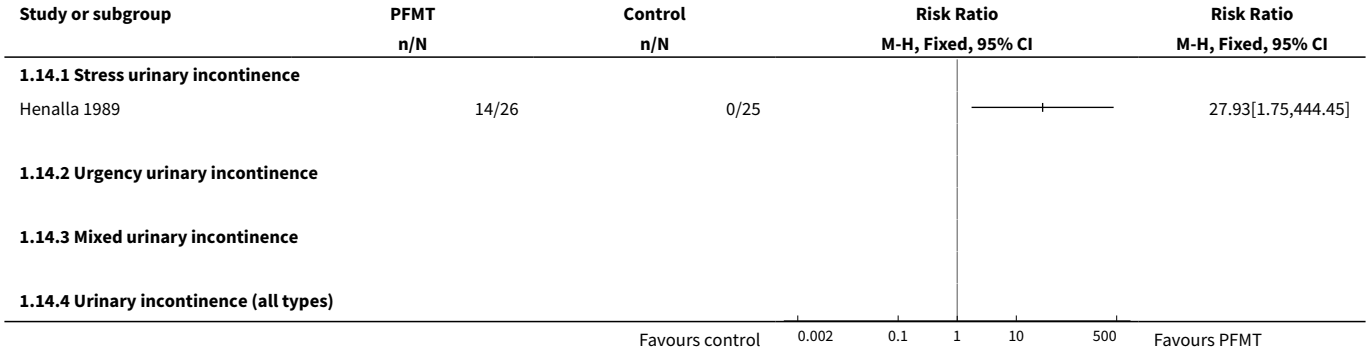
## Analysis 1.13. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 13 Participant-perceived cure at up to 1 year.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|
| **1.13.1 Stress urinary incontinence** | | | | |

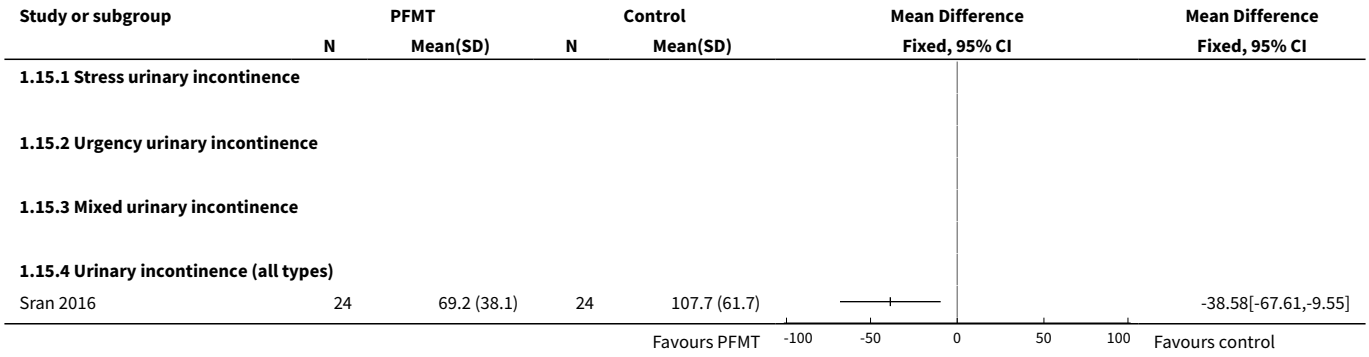
Favours control 0.001 0.1 1 10 1000 Favours PFMT




**Analysis 1.14. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 14 Participant-perceived cure or improvement at up to 1 year.**

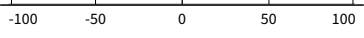



**Analysis 1.15. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 15 Urinary incontinence-specific symptom measures at 1 year (Urinary Distress Inventory long form).**



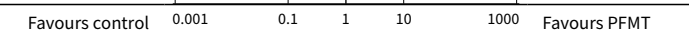
## Analysis 1.16. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 16 Urinary incontinence-specific quality of life measures at 1 year (Incontinence Impact Questionnaire long form).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|
| **1.16.1 Stress urinary incontinence** | | | | | | |
| **1.16.2 Urgency urinary incontinence** | | | | | | |
| **1.16.3 Mixed urinary incontinence** | | | | | | |
| **1.16.4 Urinary incontinence (all types)** | | | | | | |
| Sran 2016 | 24 | 20.4 (30.7) | 24 | 62.4 (98.5) | | -41.91[-83.2,-0.62] |



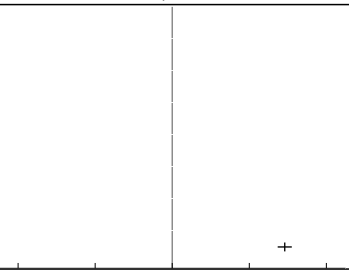

## Analysis 1.17. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 17 Participant-perceived satisfaction.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| **1.17.1 Stress urinary incontinence** | | | | | |
| Bø 1999 | 21/25 | 2/30 | | 25.91% | 12.6[3.27,48.59] |
| Castro 2008 | 15/26 | 5/24 | | 74.09% | 2.77[1.19,6.46] |
| **Subtotal (95% CI)** | **51** | **54** | | **100%** | **5.32[2.63,10.74]** |
| Total events: 36 (PFMT), 7 (Control) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=3.85, df=1(P=0.05); $I^2$=74.03% | | | | | |
| Test for overall effect: Z=4.66(P<0.0001) | | | | | |
| **1.17.2 Urgency urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.17.3 Mixed urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.17.4 Urinary incontinence (all types)** | | | | | |
| Burgio 1998 | 45/58 | 14/50 | | 100% | 2.77[1.74,4.41] |
| **Subtotal (95% CI)** | **58** | **50** | | **100%** | **2.77[1.74,4.41]** |
| Total events: 45 (PFMT), 14 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Z=4.29(P<0.0001) | | | | | |
| Test for subgroup differences: $Chi^2$=2.29, df=1 (P=0.13), $I^2$=56.37% | | | | | |



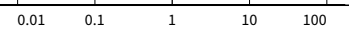
## Analysis 1.18. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 18 Perception of improvement (visual analogue scale).

| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Fixed, 95% CI | Mean Difference Fixed, 95% CI |
|---|---|---|---|---|---|---|
| **1.18.1 Stress urinary incontinence** | | | | | | |
| **1.18.2 Urgency urinary incontinence** | | | | | | |
| **1.18.3 Mixed urinary incontinence** | | | | | | |
| **1.18.4 Urinary incontinence (all types)** | | | | | | |
| Leong 2015 | 27 | 8.7 (1) | 28 | 1.4 (0.7) | | 7.3[6.84,7.76] |



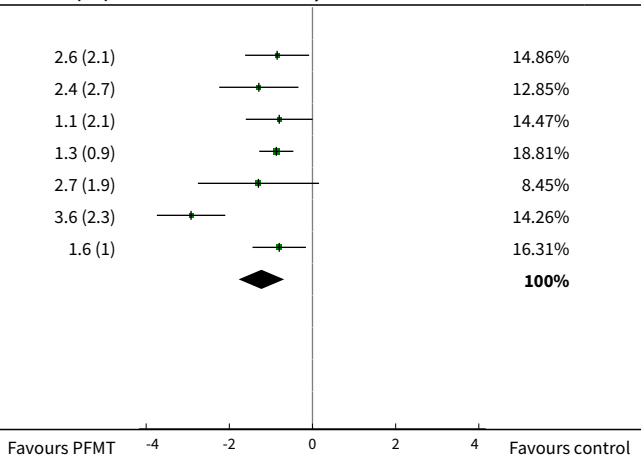

## Analysis 1.19. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 19 Number of women needing further treatment.

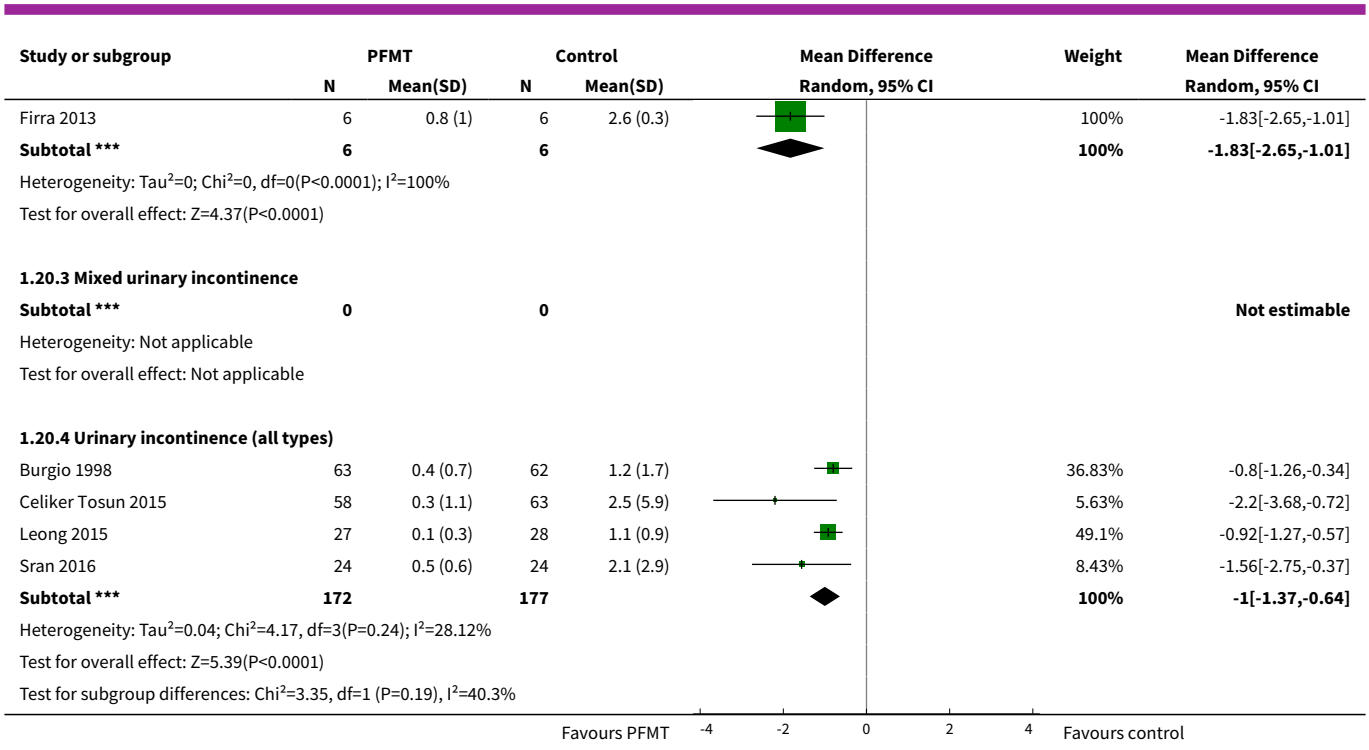
| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|
| **1.19.1 Stress urinary incontinence** | | | | |
| Bø 1999 | 4/25 | 28/30 | | 0.17[0.07,0.42] |
| **1.19.2 Urgency urinary incontinence** | | | | |
| **1.19.3 Mixed urinary incontinence** | | | | |
| **1.19.4 Urinary incontinence (all types)** | | | | |
| Burgio 1998 | 8/57 | 37/49 | | 0.19[0.1,0.36] |



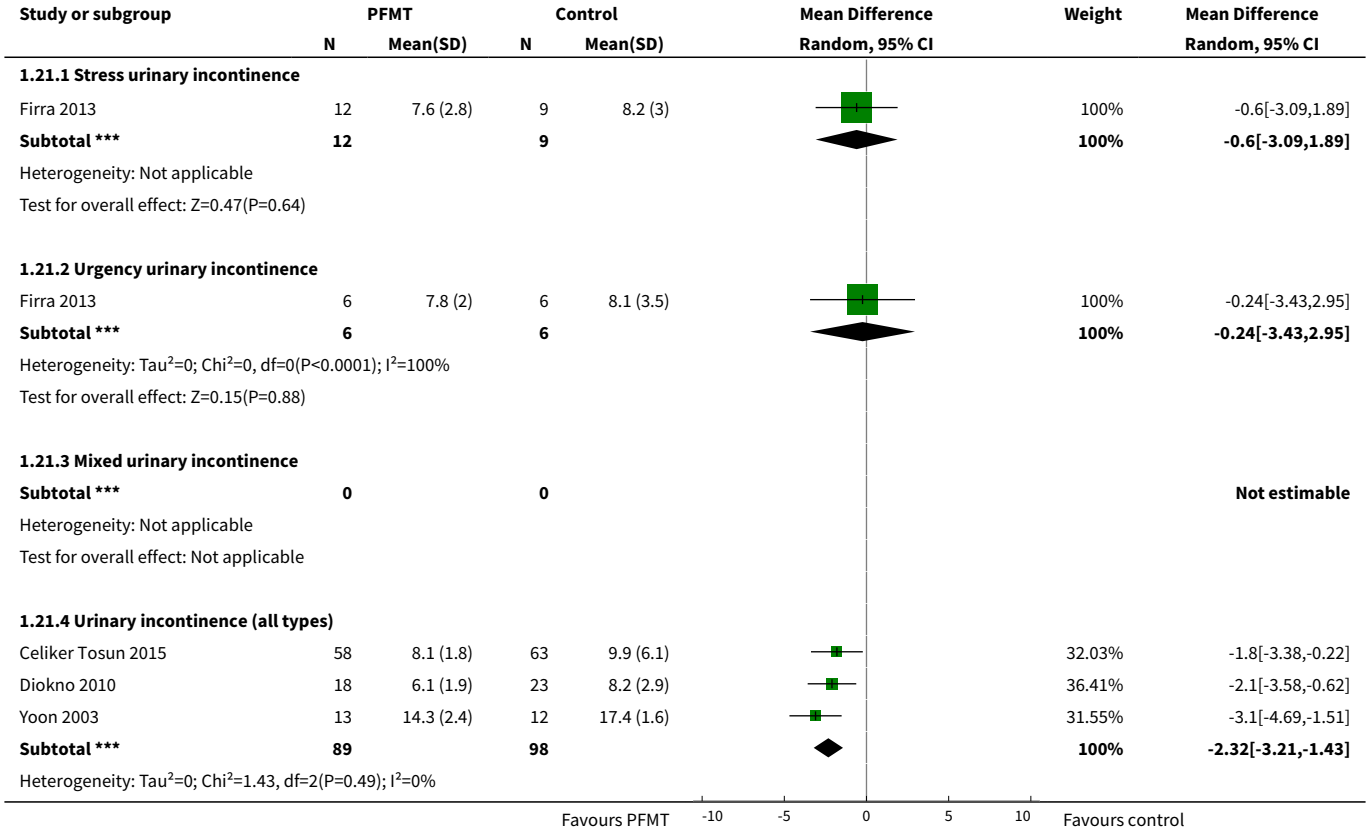

## Analysis 1.20. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 20 Number of leakage episodes in 24 hours.

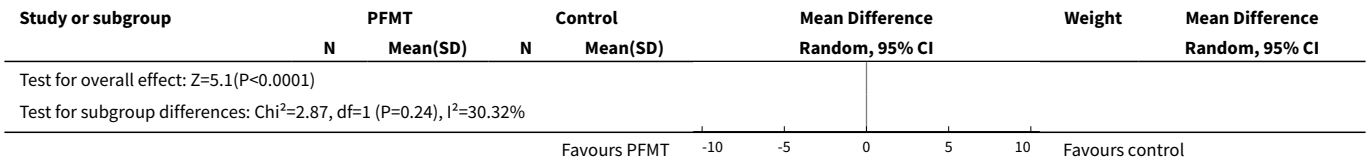
| Study or subgroup | PFMT N | PFMT Mean(SD) | Control N | Control Mean(SD) | Mean Difference Random, 95% CI | Weight | Mean Difference Random, 95% CI |
|---|---|---|---|---|---|---|---|
| **1.20.1 Stress urinary incontinence** | | | | | | | |
| Asklund 2017 | 62 | 1.7 (2.3) | 61 | 2.6 (2.1) | | 14.86% | -0.85[-1.62,-0.08] |
| Burns 1993 | 43 | 1.1 (1.4) | 39 | 2.4 (2.7) | | 12.85% | -1.29[-2.24,-0.34] |
| Bø 1999 | 25 | 0.3 (0.7) | 30 | 1.1 (2.1) | | 14.47% | -0.8[-1.6,0] |
| Castro 2008 | 26 | 0.4 (0.5) | 24 | 1.3 (0.9) | | 18.81% | -0.87[-1.28,-0.46] |
| Firra 2013 | 12 | 1.4 (1.4) | 9 | 2.7 (1.9) | | 8.45% | -1.3[-2.76,0.16] |
| Lagro-Janssen 1991a | 33 | 0.7 (0.8) | 33 | 3.6 (2.3) | | 14.26% | -2.92[-3.74,-2.1] |
| McLean 2013 | 18 | 0.8 (0.9) | 17 | 1.6 (1) | | 16.31% | -0.8[-1.44,-0.16] |
| **Subtotal ***** | **219** | | **213** | | | **100%** | **-1.23[-1.78,-0.68]** |
| Heterogeneity: $Tau^2$=0.37; $Chi^2$=22.02, df=6(P=0); $I^2$=72.75% | | | | | | | |
| Test for overall effect: Z=4.4(P<0.0001) | | | | | | | |
| **1.20.2 Urge urinary incontinence** | | | | | | | |



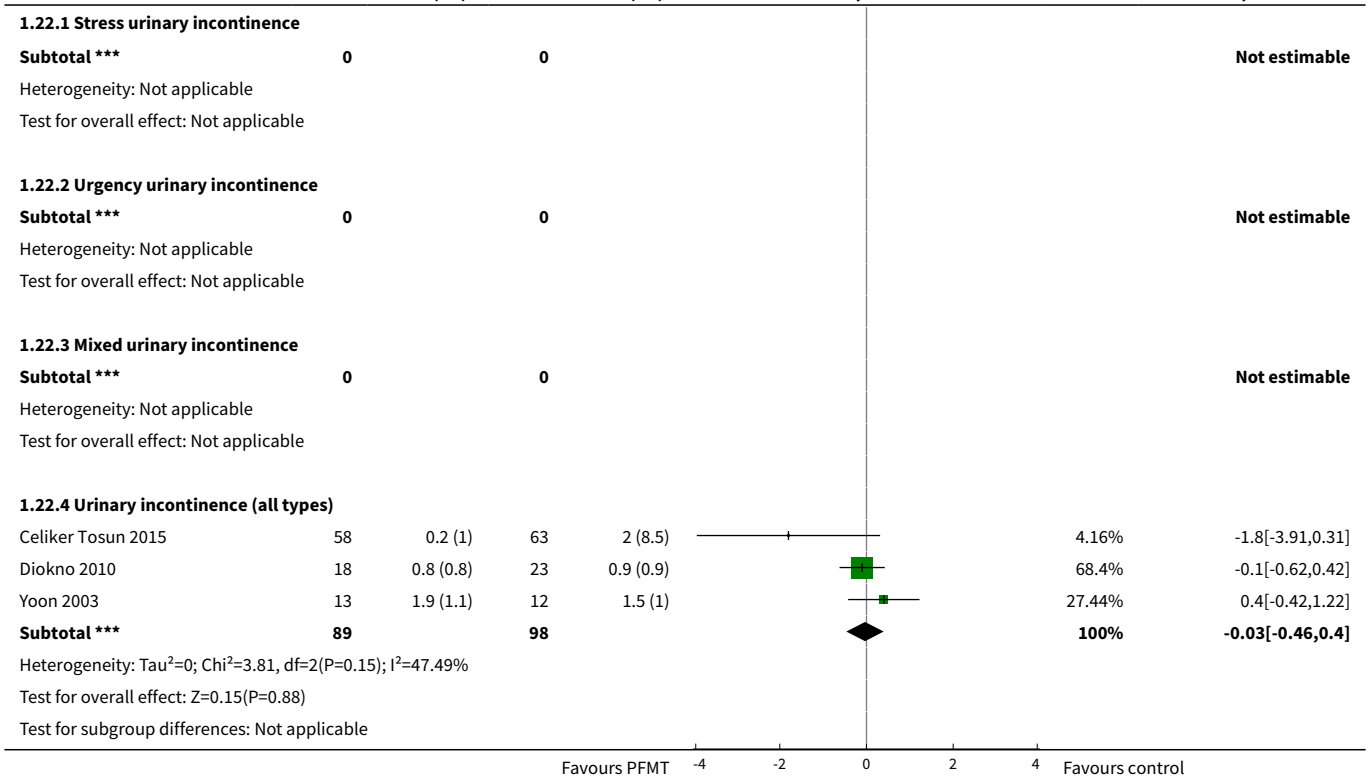



## Analysis 1.21. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 21 Number of micturitions during the day (frequency).

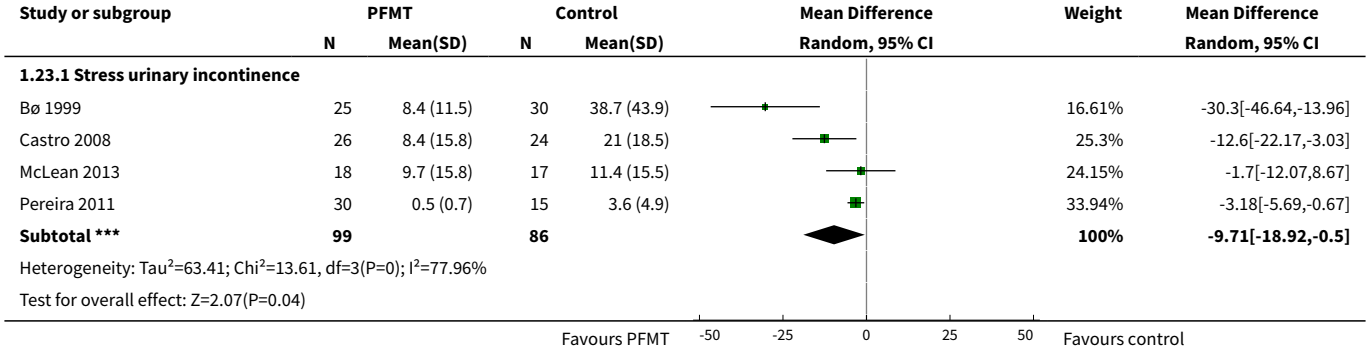


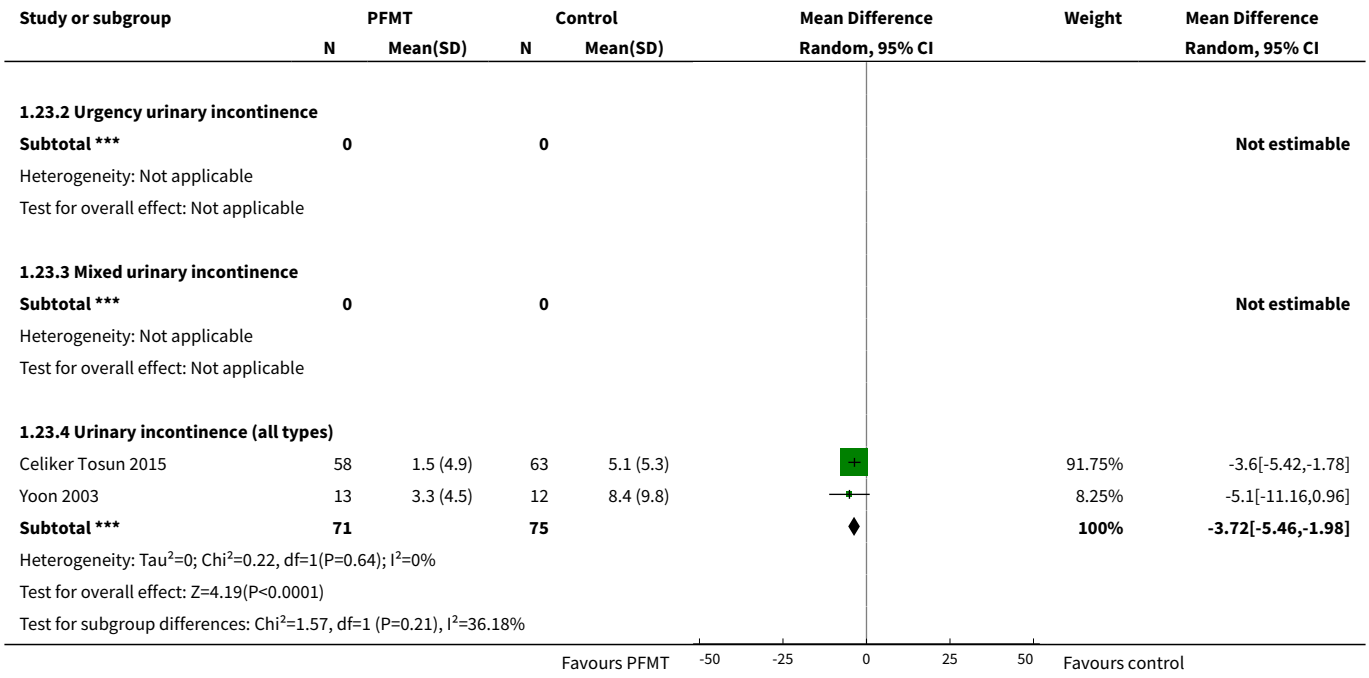



## Analysis 1.22. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 22 Number of micturitions during the night (nocturia).

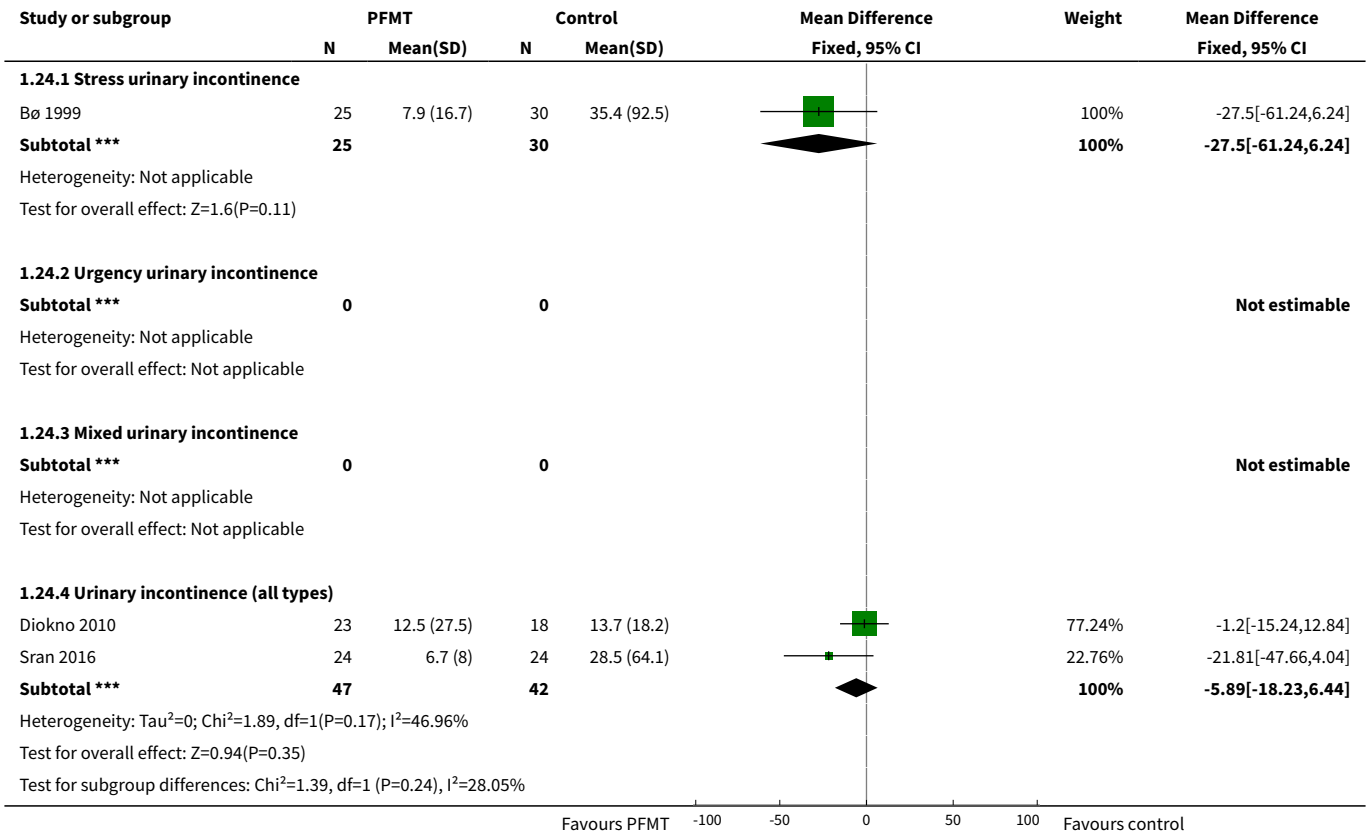



## Analysis 1.23. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 23 Short (up to 1 hour) pad test measured as grams of urine.



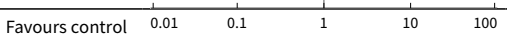



## Analysis 1.24. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 24 Long (24 hours) pad test measured as grams of urine.



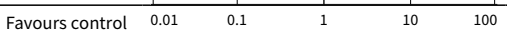
## Analysis 1.25. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 25 Number cured on short pad test (objective) after treatment.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| **1.25.1 Stress urinary incontinence** | | | | | |
| Aksac 2003 | 15/20 | 0/10 | | 14.41% | 16.24[1.07,246.51] |
| Bø 1999 | 11/25 | 2/30 | | 39.92% | 6.6[1.61,27.03] |
| Castro 2008 | 12/26 | 2/24 | | 45.67% | 5.54[1.38,22.24] |
| **Subtotal (95% CI)** | **71** | **64** | | **100%** | **7.5[2.89,19.47]** |
| Total events: 38 (PFMT), 4 (Control) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=0.52, df=2(P=0.77); $I^2$=0% | | | | | |
| Test for overall effect: Z=4.14(P<0.0001) | | | | | |
| **1.25.2 Urgency urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.25.3 Mixed urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.25.4 Urinary incontinence (all types)** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| Test for subgroup differences: Not applicable | | | | | |

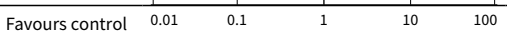



## Analysis 1.26. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 26 Number cured or improved on short pad test (objective) after treatment.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| **1.26.1 Stress urinary incontinence** | | | | | |
| Aksac 2003 | 20/20 | 2/10 | | 75.95% | 4.3[1.44,12.8] |
| Henalla 1989 | 17/26 | 0/25 | | 11.79% | 33.7[2.14,532.01] |
| Henalla 1990 | 4/8 | 0/7 | | 12.25% | 8[0.51,126.67] |
| **Subtotal (95% CI)** | **54** | **42** | | **100%** | **8.22[3.17,21.28]** |
| Total events: 41 (PFMT), 2 (Control) | | | | | |
| Heterogeneity: $Tau^2$=0; $Chi^2$=2.36, df=2(P=0.31); $I^2$=15.33% | | | | | |
| Test for overall effect: Z=4.34(P<0.0001) | | | | | |
| **1.26.2 Urgency urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |

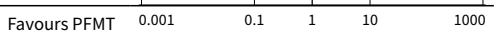


| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.26.3 Mixed urinary incontinence** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| **1.26.4 Urinary incontinence (all types)** | | | | | |
| **Subtotal (95% CI)** | **0** | **0** | | | **Not estimable** |
| Total events: 0 (PFMT), 0 (Control) | | | | | |
| Heterogeneity: Not applicable | | | | | |
| Test for overall effect: Not applicable | | | | | |
| Test for subgroup differences: Not applicable | | | | | |



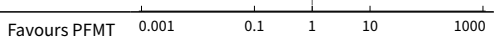

## Analysis 1.27. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 27 Number of women with sex life spoilt by urinary incontinence.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|
| **1.27.1 Stress urinary incontinence** | | | | |
| Bø 1999 | 4/20 | 13/25 | | 0.38[0.15,1] |
| **1.27.2 Urgency urinary incontinence** | | | | |
| **1.27.3 Mixed urinary incontinence** | | | | |
| **1.27.4 Urinary incontinence (all types)** | | | | |




## Analysis 1.28. Comparison 1 Pelvic floor muscle training (PFMT) versus no treatment, placebo or control, Outcome 28 Number of women with urinary incontinence during intercourse.

| Study or subgroup | PFMT n/N | Control n/N | Risk Ratio M-H, Fixed, 95% CI | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|
| **1.28.1 Stress urinary incontinence** | | | | |
| Bø 1999 | 2/20 | 10/25 | | 0.25[0.06,1.01] |
| **1.28.2 Urgency urinary incontinence** | | | | |
| **1.28.3 Mixed urinary incontinence** | | | | |
| **1.28.4 Urinary incontinence (all types)** | | | | |



# APPENDICES

## Appendix 1. Cochrane Incontinence Specialised Register search strategy

The terms used to search the Cochrane Incontinence Specialised Register are given below:

({DESIGN.CCT*} OR {DESIGN.RCT*}) AND ({INTVENT.PHYS.PFMT*} OR {INTVENT.PHYS.BIOFEED*}) AND {TOPIC.URINE.INCON*})

(All searches were of the keyword field of Reference Manager 2012).

## Appendix 2. Search strategies for the Brief Economic Commentary (BEC)

We performed additional searches for the BEC in the following databases:

- MEDLINE on OvidSP (1946 to Week 4 July 2018) was searched on 2 August 2018;
- Embase on OvidSP (1974 to Week 31 2018) was searched on 2 August 2018;
- NHS Economic Evaluation Database (NHS EED) on the Centre for Reviews and Dissemination website (www.crd.york.ac.uk/CRDWeb/) was searched on 3 August 2018.

The economic evaluation search filters used for MEDLINE and Embase are those developed by the Centre for Reviews and Dissemination to populate NHS EED and are freely available on their webpages.

The search strategies used are given below.

**MEDLINE on OvidSP** (1946 to Week 4 July 2018) was searched on 2 August 2018 using the following search strategy:

1. Economics/

2. exp "costs and cost analysis"/

3. Economics, Dental/

4. exp economics, hospital/

5. Economics, Medical/

6. Economics, Nursing/

7. Economics, Pharmaceutical/

8. (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.

9. (expenditure$ not energy).ti,ab.

10. value for money.ti,ab.

11. budget$.ti,ab.

12. or/1-11

13. ((energy or oxygen) adj cost).ti,ab.

14. (metabolic adj cost).ti,ab.

15. ((energy or oxygen) adj expenditure).ti,ab.

16. or/13-15

17. 12 not 16

18. letter.pt.

19. editorial.pt.

20. historical article.pt.

21. or/18-20

22. 17 not 21

23. exp animals/ not humans/

24. 22 not 23

25. (incontinen$ or continen$).tw.

26. exp urinary incontinence/

27. "pelvic floor"/

28. pelvic floor disorders/

29. (pelvi$ adj2 floor).tw.

30. nycturia.tw.

31. ((urin$ or bladder) adj5 sphincter$).tw.

32. ((bladder or detrusor or vesic$) adj5 (instability or stab$ or unstable or irritab$ or hyperreflexia or dys?ynerg$ or dyskinesi$ or irritat $)).tw.

33. (urethra$ adj2 sphincter$).tw.

34. (bladder adj2 neck).tw.

35. (urin$ adj2 (leak$ or urge$ or frequen$)).tw.

36. dribbl$.tw.

37. bladder, neurogenic/

38. (vesic$ adj1 (neck$ or cervi$)).tw.

39. ((bladder or detrusor or vesic$) adj2 (hyper$ or overactiv$)).tw.

40. (detrusor adj1 sphincter$).tw.

41. (spinal adj2 bladder$).tw.

42. (bladder$ adj2 (neuropath$ or neurogen$ or neurolog$)).tw.

43. (nervous adj1 (pollakisur$ or pollakiur$)).tw.

44. urinary bladder, overactive/

45. or/25-44

46. 24 and 45

47. Physical Therapy Modalities/

48. Pelvic Floor/

49. Exercise therapy/

50. Resistance training/

51. Biofeedback, psychology/

52. perineomet$.tw.

53. (pelvi$ adj5 rehab$).tw.

54. kegel*.tw.

55. (pelvi* adj4 (exercis* or train* or muscle*)).tw.

56. PFMT.tw.

57. or/47-56

58. 46 and 57

**Embase** on OvidSP (1974 to Week 31 2018) was searched on 2 August 2018 using the following search strategy:

1. Health Economics/

2. exp Economic Evaluation/

3. exp Health Care Cost/

4. pharmacoeconomics/

5. (econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.

6. (expenditure$ not energy).ti,ab.

7. (value adj2 money).ti,ab.

8. budget$.ti,ab.

9. or/1-8

10. letter.pt.

11. editorial.pt.

12. note.pt.

13. or/10-12

14. 9 not 13

15. (metabolic adj cost).ti,ab.

16. ((energy or oxygen) adj cost).ti,ab.

17. ((energy or oxygen) adj expenditure).ti,ab.

18. 15 or 16 or 17

19. 14 not 18

20. animal/

21. exp animal experiment/

22. nonhuman/

23. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).ti,ab,sh.

24. 20 or 21 or 22 or 23

25. exp human/

26. human experiment/

27. 25 or 26

28. 24 not (24 and 27)

29. 19 not 28

30. conference abstract.pt.

31. 29 not 30

32. incontinence/ or mixed incontinence/ or stress incontinence/ or urge incontinence/ or urine incontinence/

33. continence/

34. overactive bladder/

35. micturition disorder/ or lower urinary tract symptom/ or pollakisuria/

36. urinary dysfunction/ or bladder instability/ or detrusor dyssynergia/ or neurogenic bladder/ or urinary urgency/ or urine extravasation/

37. (incontinen$ or continen$).tw.

38. ((bladder or detrusor or vesic$) adj5 (instab$ or stab$ or unstab* or irritab$ or hyperreflexi$ or dys?ynerg$ or dyskinesi$ or irritat$)).tw.

39. (urin$ adj2 leak$).tw.

40. ((bladder or detrusor or vesic$) adj2 (hyper$ or overactiv$)).tw.

41. (bladder$ adj2 (neuropath$ or neurogen* or neurolog$)).tw.

42. (nervous adj pollakisur$).tw.

43. or/32-42

44. 31 and 43

45. pelvic floor muscle training/

46. exp feedback system/

47. kegel*.tw.

48. (pelvi* adj4 (exercis* or train* or muscle*)).tw.

49. PFMT.tw.

50. perineomet$.tw.

51. (pelvi$ adj5 rehab$).tw.

52. or/45-51

53. 44 and 52

**NHS Economic Evaluation Database (NHS EED)** on the Centre for Reviews and Dissemination website (www.crd.york.ac.uk/CRDWeb/) was searched on 3 August 2018 using the following search strategy:

1 (incontinen* ):TI OR (continen*):TI IN NHSEED

2 (MeSH DESCRIPTOR urinary incontinence EXPLODE ALL TREES) IN NHSEED

3 MeSH DESCRIPTOR pelvic floor EXPLODE ALL TREES IN NHSEED

4 MeSH DESCRIPTOR pelvic floor disorders EXPLODE ALL TREES IN NHSEED

5 (floor adj2 pelvi* ):TI OR (pelvi* adj2 floor):TI IN NHSEED

6 (floor adj2 pelvi* ):TI OR (pelvi* adj2 floor):TI IN NHSEED

7 (nycturia):TI IN NHSEED

8 ((urin* or bladder) adj5 sphincter*):TI OR (sphincter* adj5 (urin* or bladder)):TI IN NHSEED

9 ((bladder OR detrusor OR vesic*) ADJ5 (instability OR stab* OR unstable OR irritab* OR hyperreflexia OR dysynerg* OR dyskinesi* OR irritat*)):TI OR ((instability OR stab* OR unstable OR irritab* OR hyperreflexia OR dysynerg* OR dyskinesi* OR irritat*) ADJ5 (bladder OR detrusor OR vesic*) ):TI IN NHSEED

10 ((bladder OR detrusor OR vesic*) ADJ5 (instability OR stab* OR unstable OR irritab* OR hyperreflexia OR dyssynerg* OR dyskinesi* OR irritat*)):TI OR ((instability OR stab* OR unstable OR irritab* OR hyperreflexia OR dyssynerg* OR dyskinesi* OR irritat*) ADJ5 (bladder OR detrusor OR vesic*) ):TI IN NHSEED

11 (urethra* ADJ2 sphincter*):TI OR (sphincter* ADJ2 urethra* ):TI IN NHSEED

12 (bladder ADJ2 neck):TI OR (neck ADJ2 bladder ):TI IN NHSEED

13 (urin* ADJ2 (leak* OR urge* OR frequen*)):TI OR ((leak* OR urge* OR frequen*) ADJ2 urin* ):TI IN NHSEED

14 (dribbl*):TI IN NHSEED

15 MeSH DESCRIPTOR Urinary Bladder, Neurogenic EXPLODE ALL TREES IN NHSEED

16 (vesic* ADJ1 (neck* OR cervi*)):TI OR ((neck* OR cervi*) ADJ1 vesic*):TI IN NHSEED

17 ((bladder OR detrusor OR vesic*) ADJ2 (hyper* OR overactiv*)):TI OR ((hyper* OR overactiv*) ADJ2 (bladder OR detrusor OR vesic*)):TI IN NHSEED

18 (detrusor ADJ1 sphincter*):TI OR (sphincter* ADJ1 detrusor):TI IN NHSEED

19 (detrusor ADJ1 sphincter*):TI OR (sphincter* ADJ1 detrusor):TI IN NHSEED

20 (spinal ADJ2 bladder*):TI OR (bladder* ADJ2 spinal):TI IN NHSEED

21 (bladder* ADJ2 (neuropath* OR neurogen* OR neurolog*)):TI OR ((neuropath* OR neurogen* OR neurolog*) ADJ2 bladder*):TI IN NHSEED

22 (nervous ADJ1 (pollakisur* OR pollakiur*)):TI OR ((pollakisur* OR pollakiur*) ADJ1 nervous):TI IN NHSEED

23 MeSH DESCRIPTOR Urinary Bladder, overactive EXPLODE ALL TREES IN NHSEED

24 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23

25 MeSH DESCRIPTOR Physical Therapy Modalities EXPLODE ALL TREES IN NHSEED

26 MeSH DESCRIPTOR Pelvic Floor EXPLODE ALL TREES IN NHSEED

27 MeSH DESCRIPTOR Exercise therapy EXPLODE ALL TREES IN NHSEED

28 MeSH DESCRIPTOR Resistance training EXPLODE ALL TREES IN NHSEED

29 MeSH DESCRIPTOR Biofeedback, psychology EXPLODE ALL TREES IN NHSEED

30 (perineomet*):TI IN NHSEED

31 (pelvi* ADJ5 rehab*):TI OR (rehab* ADJ5 pelvi*):TI IN NHSEED

32 (kegel*):TI IN NHSEED

33 (kegel*):TI IN NHSEED

34 (pelvi* ADJ4 (exercis* OR train* OR muscle*)):TI OR ((exercis* OR train* OR muscle*) ADJ4 pelvi*):TI IN NHSEED

35 (PFMT):TI IN NHSEED

36 #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35

37 #24 AND #36

## Appendix 3. PFMT protocols

| Study ID | VPFMC taught/ confirmed | Description | Total VPFMC | Duration of pro-gramme | Supervi-sion |
|---|---|---|---|---|---|

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | | | **per day** | | |
| Aksac 2003 | Taught by: therapist<br><br>Confirmed by: vaginal palpation, while keeping abdominal and buttock muscles relaxed | Number of VPFMC per set: 10<br>Duration of hold: 5 seconds<br>Duration of rest: 10 seconds<br>Number sets per day: 3<br>Body position(s): not reported<br>Type(s) of contraction: sustained<br>Other exercise(s): contractions progressed at 2 weeks to 10 seconds' hold and 20 seconds' rest, home treatment<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | 30 | 8 weeks | Weekly clinic visits |
| Asklund 2017 | Taught by: smartphone app<br><br>Confirmed by: basic contraction for the participant to identify the correct pelvic floor muscle contraction (not face-to-face) | Number of VPFMC per set: progressive (from 8 to 62)<br>Duration of hold: 2–59<br>Duration of rest: 2–59<br>Number sets per day: 3<br>Body position(s): standing, lifting, walking<br>Type(s) of contraction: exercises to find the right muscle + strength, endurance, quick contraction and knack<br>Other exercise(s): information about SUI, PFMT exercises at different difficulty levels (6 basic and 6 advanced) with graphic support; lifestyle recommendations<br>Adherence strategy(s): possibility to set 3 reminders/day in the app; each exercise description included graphics showing the duration and intensity of each contraction/relaxation<br>Adherence measures: statistics table for exercise diary in the app; those who used this function (n = 52/61) registered a mean of 141 exercises/person during the study, which would correspond to a mean of 1.6 exercises/day during a 90-day period | progressive (from 24 to 186) | 3 months | no face-to-face contact with the participants during the study |
| Bertotto 2017 | Taught by: physiotherapist<br><br>Comfirmed by: digital palpation | Number of VPFMC per set: 24–30<br>Duration of hold: 6–10<br>Duration of rest: 4–10<br>Number sets per day: 1–3<br>Body position(s): supine lying, sitting and standing | Progressive from 24 to 70 | 4 weeks | 20 minutes, twice per week clinic visits |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | | Type(s) of contraction: sustained, phasic lasting 2 sec, phasic lasting 3–5 sec, guided imagery training<br>Other exercise(s): none<br>Adherence strategy(s): none<br>Adherence measures: none | | | |
| Beutten-muller 2010 | Taught by: physical therapist<br>Confirmed by: not reported, but assessed by the evaluator prior to treatment | Number of VPFMC per set: 8<br>Duration of hold: 5 seconds<br>Duration of rest: not reported<br>Number sets per day: not reported<br>Body position(s): supine with knee bent, sitting on a chair or gym ball, on all fours, and standing<br>Type(s) of contraction: submaximal, maximal/long and short contractions<br>Other exercise(s): proprioceptive exercises such as sitting and hopping around a ball, movements that raise the pelvis (e.g. anteversion, retroversion, lateralisation and circumduction)<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | Not reported | 6 weeks | 20-minute twice per week clinic visits<br>* Except during menstruation or due to other complications |
| Bidmead 2002 | Taught by: physical therapist<br>Confirmed by: not reported | Number of VPFMC per set: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Number sets per day: not reported<br>Body position: not reported<br>Type(s) of contraction: not reported<br>Other treatment(s): not reported<br>Adherence strategy(s): none reported<br>Adherence measure: exercise diary; compliance with PFM exercises was generally good with 75% of participants performing the exercises > 3 times per week | Not reported | 14 weeks | 5 clinic visits over 14-week period (weeks 1, 3, 6, 10 and 14) |
| Burgio 1998 | Taught by: nurse practitioner<br>Confirmed by: anorectal biofeedback while | Number of VPFMC per set: 15<br>Duration of hold: based on each participant's ability and gradually increased across multiple sessions to a maximum of 10 seconds<br>Duration of rest: based on each participant's ability | 45 | 8 weeks | 4 clinic visits at 2-week intervals |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | keeping abdominal muscles relaxed | Number sets per day: 3<br>Body position(s): supine, sitting, standing<br>Type(s) of contraction: not reported<br>Other treatment(s): Knack and interrupting or slowing urine stream once per day<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | | | |
| Burns 1993 | Taught by: nurse trained in biofeedback techniques<br>Confirmed by: biofeedback to teach the participant to relax and contract the pelvic muscles | Number of VPFMC per set: 10 quick and 10 sustained (× 2 sets)<br>Duration of hold: 10 contractions held for 3 seconds and 10 contractions held for 10 seconds<br>Duration of rest: not reported<br>Number sets per day: 4<br>Body position(s): not reported<br>Type(s) of contraction: fast, sustained<br>Other treatment(s): videotape describing exercise protocol for home exercises<br>Intervention progressed 20 per set to a daily maximum of 200<br>Adherence strategy(s): weekly and post-treatment 3- and 6-month telephone reminder calls for the appointments; weekly home exercise reminder cards mailed between visits<br>Adherence measures: exercise diary; results not reported | Progressive from 80 to 200 | 8 weeks | Weekly clinic visits |
| Bø 1999 | Taught by: physical therapist<br>Confirmed by: vaginal palpation | Number of VPFMC per set: 8–12 high-intensity (close to maximal) with 3–4 fast contractions added at the end of each hold; if ability to hold the contraction improved the duty cycle was progressed each month<br>Duration of hold: 6–8 seconds for the high intensity contractions<br>Duration of rest: 6 seconds<br>Number sets per day: 3<br>Body position(s): supine, kneeling, sitting, standing; all with legs apart. Participants used preferred position<br>Type(s) of contraction: sustained high-intensity contractions and quick contractions<br>Other treatment(s): verbal information on the PFM and lower urinary tract anatomy and physiology and on continence mechanisms | 36 | 6 months | 45-minute per week exercise class<br>Monthly clinic visit with physical therapist |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | | Body awareness, breathing, relaxation exercises and strength training exercises for the back, abdominal and thigh muscles<br>Adherence strategy(s): audiotape with verbal guidance for home training<br>Adherence measures: exercise diary; mean adherence with treatment was 93% (SE 1.5%) for PFMT | | | |
| Carneiro 2010 | Taught by: physical therapist<br>Confirmed by: vaginal palpation | Number of VPFMC per set: 8–12 (5 sets total)<br>Duration of hold: 6–10 seconds<br>Duration of rest: not reported<br>Number sets per day: 5<br>Body position(s): sitting, standing<br>Type(s) of contraction: sustained<br>Other treatment(s): verbal information about PFM function and visualisation of pelvic floor components using anatomical figures<br>5 minutes of proprioceptive exercises sitting on a 75-cm diameter therapeutic ball<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | Progressive from 40 to 60 | 8 weeks | 30-minute, twice per week clinic visits |
| Castro 2008 | Taught by: physical therapist<br>Confirmed by: vaginal palpation | Number of VPFMC and duration of hold and rest:<br>• 5 contractions held 10 seconds with 10-second recovery<br>• 10 contractions held 5 seconds with 5-second recovery<br>• 20 contractions held 2 seconds with 2-second recovery<br>• 20 contractions held 1 second with 1-second recovery<br>• 5 contractions with cough<br>Number sets per day: once, 3 times per week<br>Body position(s): not reported<br>Type(s) of contraction: sustained and quick contractions<br>Other treatment(s): verbal information on the PFM and lower urinary tract anatomy and physiology and on continence mechanisms<br>Warm-up exercises for the joints and stretching exercises targeting the hip, adductor, hamstring and paravertebral muscles<br>Adherence strategy(s): not reported<br>Adherence measures: exercise diary updated by the physical therapist during each clinic visit; mean compliance of 92% in the PFMT group | 60 | 6 months | 3 group sessions per week |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| Celiker Tosun 2015 | Taught by: physiotherapist<br><br>Confirmed by: by digital palpation | Number of VPFMC per set: 12<br>Duration of hold: according to Laycock PERFECT scheme<br>Duration of rest: according to Laycock PERFECT scheme<br>Number sets per day: 3<br>Body position(s): not reported<br>Type(s) of contraction: progressive training with fast, slow and sustained contractions + knack<br>Other treatment(s): none<br>Adherence strategy(s): training diary (participants required to keep a training diary to maintain their motivation)<br>Adherence measures: not reported | 36 | 12 weeks + additional programme for those who did not reach the PFM strength goal (grade 5 on Oxford scale)) | 30 minutes 3 times per week for the first 2 weeks followed by weekly visits |
| Diokno 2010 | Taught by: urology nurse<br><br>Confirmed by: not reported | Number of VPFMC per set: 25 (5 short and 20 long contractions) and, when needed, the Knack (sneezing)<br>Duration of hold: long contractions held up to 6 seconds<br>Duration of rest: not reported<br>Number sets per day: 2<br>Body position(s): not reported<br>Type(s) of contraction: short and long contractions<br>Other treatment(s): 2-hour Microsoft PowerPoint presentation, BMP lecture with printed handouts on the lower urinary tract anatomy, the mechanism of urinary bladder function, and UI<br>Bladder training tips, if needed<br>Knack, if needed<br>Audiotape for daily use<br>Adherence strategy(s): 2–4 weeks' follow-up, including a vaginal examination if needed, measurement of PFM strength and an ability test<br>Adherence measures: not reported | 50 | 6–8 weeks | 1 teaching session<br><br>1 follow-up session after 2–4 weeks with a vaginal examination if needed and a written test on new knowledge acquired |
| Ferreira 2014 | Taught by: not reported<br><br>Confirmed by: unclear | Number of VPFMC per set: 30 sustained contractions + 4 × 30 fast contractions<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Number sets per day: 1 | 150 | 3 months | Weekly visits were made at the club during the study period to ensure motivation and ad- |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | | Body position(s): not reported<br>Type(s) of contraction: sustained, fast<br>Other treatment(s): none<br>Adherence strategy(s): none<br>Adherence measures: not reported | | | herence to PFMT both at home and after the training sessions (volleyball practice). |
| Firra 2013 | Taught by: physiotherapist<br>Confirmed by: yes, by digital examination | Number of VPFMC per set: progressive up to 23 contractions<br>Duration of hold: 10 s for sustained contraction, 2–3s for short contractions<br>Duration of rest: 20 s for sustained contraction, 6 s for short contractions<br>Number sets per day: 5<br>Body position(s): for all exercises position progressed from hook lying to sitting, standing squatting as able<br>Type(s) of contraction: sustained, maximal (short contractions) and submaximal (controlled)<br>Other treatment(s): none<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | Up to 115 | 8 weeks | 16 sessions (2 per week) |
| Henalla 1989 | Taught by: physical therapist<br>Confirmed by: vaginal palpation | Number of VPFMC per set: 5<br>Duration of hold: 5 seconds<br>Duration of rest: not reported<br>Number sets per day: 1 set per hour during the day<br>Body position(s): not reported<br>Type(s) of contraction: not reported<br>Other treatment(s): not reported<br>Adherence strategy(s): not reported<br>Adherence measure: not reported | ~ 80 | 12 weeks | Weekly clinic visit |
| Henalla 1990 | Taught by: physical therapist<br>Confirmed by: not reported | Number of VPFMC per set: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Number sets per day: not reported | Not reported | 6 weeks | Not reported |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | | Body position(s): not reported<br>Type(s) of contraction: not reported<br>Other treatment(s): not reported<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | | | |
| Hofbauer 1990 | Taught by: physical therapist<br>Confirmed by: not reported | Number of VPFMC per set: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Number sets per day: not reported<br>Body position(s): not reported<br>Type(s) of contraction: not reported<br>Other treatment(s): abdominal wall and adductor exercises and home training<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | ?? | 6 months | 20-minute twice per week clinic visits |
| Kargar Jahromi 2013 | Taught by: not clear<br>Confirmed by: not clear how/if it was confirmed | Number of VPFMC per set: 8–12 sustained + 4 rapid contractions in each position<br>Duration of hold: 6–8 sec<br>Duration of rest: 6 sec<br>Number sets per day: 3<br>Body position(s): lying, sitting, standing<br>Type(s) of contraction:<br>sustained maximal and fast contractions<br>Other treatment(s): education, body awareness and breathing<br>Adherence strategy(s): none<br>Adherence measures: not reported | 36 high intensity (close to maximum) contractions + 12 rapid contractions | 8 weeks | Training in groups once per week for 45 minute |
| Kim 2007 | Taught by: nurse<br>Confirmed by: participants were trained to exert force only on the | During the 12 weeks' intervention:<br>Number of VPFMC per set: 10 (× 2 sets)<br>Duration of hold: 10 contractions held 3 seconds and 10 additional contractions held 6–8 seconds<br>Duration of rest: 10 seconds | ~ 30 | 12 weeks | Exercise class, twice per week |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | PFM but did not give detail on how it was done | Number sets per day: twice per week<br>Body position(s): sitting, supine and standing positions with the legs apart<br>Type(s) of contraction: fast and sustained contractions<br>Other treatment(s): body awareness, breathing, and relaxation exercises. Strength training for the thigh, abdominal, and back muscles (i.e. bending the knees, tilting the pelvis backward and forward, lifting the buttocks on the back with the knees bent, raising 1 leg while lying on the back)<br>Exercises using 2 types of training balls<br>Adherence strategy(s): home training reinforced through a pamphlet illustrating PFM and strengthening exercises and a record-keeping sheet<br>Adherence measures: attendance rate to exercise sessions and exercise diary for the follow-up. Attendance ranged from 71.9% to 93.8%, with a mean of 82.4%. Exercise frequency during follow-up was every day in 30.3% of the participants, 2–3 times per week in 45.5%, and once or less per week in 24.2%<br>During 1-year follow-up:<br>Number of VPFMC per set: 13<br>Sets per day: 2–3 sets at least twice per week | | | |
| Kim 2011a | Taught by: nurse<br>Confirmed by: participants were trained to exert force on just the PFMs, but details on how this was done were lacking | Number of VPFMC per set: 10 fast and 10 sustained contractions<br>Duration of hold: 3 seconds for fast contractions, 6–8 seconds for sustained contractions<br>Duration of rest: 5 seconds for fast contractions, 10 seconds for sustained contractions<br>Number sets per day: 3<br>Body position(s): PFM contractions, without excessively straining the abdomen, performed in supine, sitting, and standing positions with legs apart<br>Type(s) of contraction: fast and sustained contractions<br>Other treatment(s): warm-up and stretching exercises 10–15 minutes; thigh and abdominal muscle strength training exercises between PFM training, and weight bearing and ball exercises<br>Home exercises 2–3 sets (PFM + 13 other exercises) at least 3 times per week (duration: approximately 30 minutes)<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | 60 | 12 weeks | 1-hour, twice per week group sessions |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| Kim 2011b | Taught by: nurse<br><br>Confirmed by: participants were trained to exert force on just the PFMs, but details on how this was done were lacking | Number of VPFMC per set: 10 fast and 10 sustained contractions<br><br>Duration of hold: 3 seconds for fast contractions, 6–8 seconds for sustained contractions<br><br>Duration of rest: 5 seconds for fast contractions, 10 seconds for sustained contractions<br><br>Number sets per day: 3<br><br>Body position(s): PFM contraction without excessively straining the abdomen, performed in supine, sitting, and standing positions with legs apart<br><br>Type(s) of contraction: fast and sustained contractions<br><br>Other treatment (s): warm-up and stretching exercise for 10–15 minutes. Strength training of the thigh and abdominal muscles, back, legs, trunk and use of an exercise ball<br><br>Adherence strategy(s): not reported<br><br>Adherence measures: attendance to exercise session and exercise diary for the follow-up. The attendance rate ranged from 63.5% to 81.1%, with a mean of 70.3%. The exercise frequency during the follow-up was reported to be every day in 35.7% of the participants, 2–3 times per week in 42.9%, and once or less per week in 21.4%. The mean exercise time was 29.3 minutes, and the mean number of contractions of the PFM was 52 times/day<br><br>Follow-up: after the 12 weeks' intervention, participants attended a 1-hour exercise class once a month for 7 months and continued a home-based programme (2–3 sets of PFM + 13 other exercises taught during the intervention) | 60 | 12 weeks | 1-hour, twice per week group sessions |
| Lagro-Janssen 1991a | Taught by: general practitioner<br><br>Confirmed by: vaginal palpation | Number of VPFMC per set: 10<br><br>Duration of hold: 6 seconds<br><br>Duration of rest: not reported<br><br>Number sets per day: 5–10<br><br>Body position(s): not reported<br><br>Type(s) of contraction: not reported<br><br>Other treatment(s): verbal information on PFMs<br><br>Adherence strategy(s): not reported<br><br>Adherence measures: participants were asked how many exercises per day they completed and how well they complied with the exercise programme (exercise diary) | 50–100 | 12 weeks | No supervision, the participants received written instructions for home practice |
| Leong 2015 | Taught by: physiotherapist | Number of VPFMC per set: 5–30 submaximal and 5–10 rapid progressive | 15–90 submaximal con- | 12 weeks | 8 physiotherapy sessions, |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | Confirmed by: vaginal digital palpation | Duration of hold: 5–10 seconds<br>progressive<br>Duration of rest: 10 seconds<br>Number sets per day: 3<br>Body position(s): lying, sitting, standing<br>Type(s) of contraction: slow submaximal (5–10 sec), maximal – fast (10 ×), Knack<br>Other treatment(s): bladder training: strategies to increase the time interval between voids by a combination of progressive void schedules, urge suppression, distraction, self-monitoring and reinforcement<br>Adherence strategy(s): none<br>Adherence measures: proportion of sessions attended (attendance rate), frequency of home exercises executed (exercise diary); mean attendance rate in the intervention group was 97.7% (SD 5.0%) and mean exercise compliance was 99.4% (SD 1.9%) | tractions + 15–30 fast contractions | | 30 minutes<br>1 per week for the first 4 weeks<br>Once per 2 weeks for the 5–12 weeks |
| McLean 2013 | Taught by: physiotherapist<br>Confirmed by: vaginal digital palpation | Number of VPFMC per set: 10 + specific exercises focused on holding or relaxing after the contraction<br>Duration of hold: ramp up over 2–4 s, hold 1–2 s<br>Duration of rest: not specified<br>Number sets per day: 3 sets<br>Body position(s): not reported<br>Type(s) of contraction: maximal voluntary contraction, Knack<br>Other treatment(s): none reported<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | 30 | 12 weeks | Weekly sessions of 30 minutes |
| Miller 1998 | Taught by: nurse<br>Confirmed by: vaginal palpation | Number of VPFMC per set: not reported<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Number sets per day: not reported<br>Body position(s): not reported<br>Type(s) of contraction: co-ordination | Not reported | 1 week | No supervision |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | | Other treatment(s): verbal information on PFM physiology and functional properties; participants were taught to practice the Knack<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | | | |
| Pereira 2011 | Taught by: physical therapist<br>Confirmed by: vaginal palpation and instructed not to use compensatory muscles | **For Group and individual PFMT intervention**<br>Number of VPFMC per set: on average, 100 contractions were performed,<br>Duration of hold: 5–10 seconds<br>Duration of rest: 10–20 seconds<br>Number sets per day: not reported<br>Body position(s): supine, sitting and standing positions<br>Type(s) of contraction: phasic and tonic contractions<br>Other treatment(s): verbal information on the PFM anatomy and continence mechanisms; the degree of difficulty progressed according to the positions adopted, the number of repetitions, and the time of sustained contractions<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | 100 | 6 weeks | 2 × 1-hour weekly sessions in clinic |
| Sar 2009 | Taught by: nurse<br>Confirmed by: vaginal palpation | Number of VPFMC per set: 30<br>Duration of hold: 1–10 seconds<br>Duration of rest: same as contraction time<br>Number sets per day: 3<br>Body position(s): supine, sitting and standing<br>Type(s) of contraction: quick flicks (1–2 second contractions), sustained progressive (5–10 seconds) contractions<br>Other treatment(s): verbal information on the PFM and lower urinary tract anatomy and physiology and on continence mechanisms; Knack<br>Adherence strategy(s): weekly telephone call to encourage exercises practice and answer questions<br>Adherence measures: not reported | 90 | 6 weeks | Weekly telephone call by the nurse |
| Solberg 2016 | Taught by: physiotherapist | Number of VPFMC per set: not reported<br>Duration of hold: not reported | 10 minutes | 12 weeks | 12 weekly sessions of 25 min of PFMT + 20 min |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | Confirmed by: vaginal digital palpation | Duration of rest: not reported<br>Number sets per day: 10 minutes<br>Body position(s): not reported<br>Type(s) of contraction: not reported<br>Other treatment(s): general exercises<br>Adherence strategy(s): not reported<br>Adherence measures: attendance rate and exercise diary; 6 women who completed PFMT, 4 attended 12 sessions of PFMT, 1 attended 11 and 1 attended 10 sessions; all completed the training diary of 10 min PFMT at home every day | | | of general exercise |
| Sran 2016 | Taught by: physiotherapist<br>Confirmed by: vaginal digital palpation | Number of VPFMC per set: variable and progressive<br>Duration of hold: according to Laycock's perfect scheme<br>Duration of rest: same as above<br>Number sets per day: 1<br>Body position(s): lying to standing<br>Type(s) of contraction:<br>progressive<br>• fast MVC<br>• slow MVC<br>• Knack (static dynamic)<br>• PFM + TA lying and standing<br>• urge suppression techniques<br>• crown (100-50-70-100)<br>• PFME walking, lunging, squatting<br>Other treatment(s): EMG biofeedback to monitor PFM EMG activity<br>Adherence strategy(s): none reported<br>Adherence measures: number of session attended (attendance rate), number of PFME done at home (exercise diary); 58% (14/24) attended all 12 sessions and 33% (8/24) 10 or 11 sessions; 33% (8/24) completed 100% of the home exercises, 33% (8/24) 70%–99%, and 1 participant 50%; 21% (5/24) of the participants did not complete, either partially or at all, their exercise diary, although most of them reported doing the exercises some of the time | 10 × 10-s MVC holds and 30 fast contractions in each position | 12 weeks | 1 hour per week × 1 week + 30 min per week × 11 weeks |
| Wells 1999 | Taught by: nurse practitioner | Number of VPFMC per set: 80<br>Duration of hold: 10 seconds | 80 | 5 months | Monthly visits for observation, coaching |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | Confirmed by: able to contract PFM was confirmed through a physical examination | Duration of rest: 10 seconds<br>Number sets per day: 1 set during the day<br>Body position(s): not reported<br>Type(s) of contraction: sustained<br>Other treatment(s): not reported<br>Adherence strategy(s): not reported<br>Adherence measures: training diary, results not reported | | | and encouragement |
| Yoon 2003 | Taught by: nurse<br>Confirmed by: weekly surface EMG biofeedback | Number of VPFMC per set: 30 strength and endurance VPFMC per day (unclear if this is 30 for both combined or 30 per type of exercise; i.e. 60), approximately 15–20 minutes per day<br>Duration of hold: not reported<br>Duration of rest: not reported<br>Number sets per day: not reported<br>Body position(s): not reported<br>Type(s) of contraction: strength (burst of intense activity lasting a few seconds) and endurance (6-second hold progressed by 1-second per week to 12 seconds)<br>Other treatment(s): not reported<br>Adherence strategy(s): not reported<br>Adherence measures: not reported | Not clear if 30 or 60 | 8 weeks | Weekly clinic visit with nurse |

BMP: behavioural modification programme; EMG: electromyography; min: minute; MVC: maximal voluntary contraction; n: number of participants; PFM: pelvic floor muscle; PFMT: pelvic floor muscle training; SD: standard deviation; SE: standard error; sec: second; SUI: stress urinary incontinence; TA: transabdominal; UI: urinary incontinence; VPFMC: voluntary pelvic floor muscle contraction.

## Appendix 4. Other urinary incontinence-specific quality of life outcomes

| Study ID | Outcome | Measure | Subscale | PFMT | Control | Difference |
|---|---|---|---|---|---|---|
| Bø 1999 | **Bristol Female Lower Urinary Tract Symptoms (BFLUTS) Questionnaire**<br>For analysis, positive findings ('a little,' 'somewhat' and 'a lot,' or 'a bit of a problem,' 'quite a problem' and 'a serious problem') were regrouped and reported as frequencies. Only the lifestyle (28–31, 33) and sex-life questions (21–24) were reported. | Number and % | Avoiding places and situations | n = 25<br>7 | n = 30<br>10 | RR 0.84, 95% CI 0.37 to 1.88 |
| | | | Interference with social life | n = 25<br>1 | n = 30<br>12 | RR 0.10, 95% CI 0.01 to 0.72 |

(Continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | Interference with physical activity | n = 25<br>11 | n = 30<br>24 | RR 0.55, 95%<br>CI 0.34 to 0.89 |
| | | | Overall interference with life | n = 25<br>14 | n = 30<br>25 | RR 0.67, 95%<br>CI 0.46 to 0.99 |
| | | | Unsatisfied if had to spend rest of life as now | n = 25<br>10 | n = 30<br>11 | RR 0.11, 95%<br>CI 0.02 to 0.79 |
| | | | Sex-life spoilt by urinary symptoms | n = 20<br>3 | n = 25<br>13 | RR 0.29, 95%<br>CI 0.10 to 0.87 |
| | | | Problem with sex-life being spoilt | n = 20<br>2 | n = 25<br>13 | RR 0.19, 95%<br>CI 0.05 to 0.76 |
| | | | Problem with painful intercourse | n = 20<br>2 | n = 25<br>10 | RR 0.25, 95%<br>CI 0.06 to 1.01 |
| | | | Urinary incontinence with intercourse | n = 20<br>2 | n = 25<br>10 | RR 0.25, 95%<br>CI 0.06 to 1.01 |
| | **Social Activity Index**<br><br>Provides a summation of scores for a visual analogue scale for perception of difficulty participating in 9 specified social situations. A lower score indicates problem is perceived to be greater. This index has established reproducibility in women with SUI (Bø 1994). | Mean score (SD) | NA | n = 25<br>9.3 (1.0) | n = 30<br>7.9 (2.2) | MD 1.4, 95%<br>CI 0.4 to 2.4 |
| Diokno 2010 | **Sandvik's Incontinence Severity Index (ISI) for Female Urinary Incontinence** (3-point scale)<br><br>Questions assess the degree of UI:<br><br>Frequency: 1. How often do you experience urinary leakage? Scale: 1 = less than once a month, 2 = a few times a month, 3 = a few times per week, 4 = every day or night (or both).<br><br>Quantity: 2. How much urine do you lose each time? Scale: 1 = drops, 2 = small splashes and 3 = more. Note: on the 3-level severity index, responses to this question are aggregated into drops (1) or more (2). | Number and % | — | n = 23 | n = 18 | — |
| | | | Slight | 13 (56.5%) | 5 (22.2%) | RR 2.03, 95%<br>CI 0.89 to 4.65 |
| | | | Moderate | 5 (21.7%) | 7 (38.9%) | RR 0.78, 95%<br>CI 0.27 to 2.29 |
| | | | Severe | 5 (21.7%) | 7 (38.9%) | RR 0.78, 95%<br>CI 0.27 to 2.29 |

(Continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| | The Severity Index is created by multiplying the result of questions 1 (quantity) and 2 (frequency), resulting in the following index values whereby 1–2 = slight, 3–4 = moderate and 6–8 = severe<br><br>This index has been validated in women with UI (Grade C, Kelleher 2013; Sandvik 2000). | | | | | |
| Firra 2013 | **York Incontinence Perception Scale (YIPS)**<br><br>This is a self-administered questionnaire that measures self-perceived limitations in social settings and the ability to cope with UI in a constructive way (Grade B, Kelleher 2013). Scores on the scale range from 8 to 56, with the higher score indicating improved coping skills (Lee 1995). | Mean SUI score (SD) | NA | n = 12<br><br>44.8 (6.3) | n = 9<br><br>29.9 (2.2) | MD 14.90, 95% CI 11.06 to 18.74 |
| | | Mean UUI score (SD) | NA | n = 6<br><br>47.0<br><br>(5.5) | n = 6<br><br>28.8<br><br>(2.9) | MD 18.20, 95% CI 13.22 to 23.18 |

CI: confidence interval; MD: mean difference; n: number of participants; NA: not applicable; PFMT: pelvic floor muscle training; RR: risk ratio; SD: standard deviation; SUI: stress urinary incontinence; UI: urinary incontinence.

## Appendix 5. Other leakage outcomes

| Study ID | Outcome | Measure | PFMT | Control | Difference |
|---|---|---|---|---|---|
| Bø 1999 | **Leakage index**<br><br>Perceived frequency of leakage with 7 prespecified types of exertion. Higher score indicates more perceived leakage. | Mean (SD) | n = 25<br><br>1.9 (0.5) | n = 30<br><br>3.1 (0.6) | MD –1.2, 95%<br><br>CI –1.5 to –0.9 |
| Celiker Tosun 2015 | **Urgent voiding on urinary diary** | Mean<br><br>(SD) | n = 58<br><br>0.5 (1.8) | n = 63<br><br>5.0 (3.2) | MD –3.90, 95% CI –5.22 to –2.58 |
| | **Stop test** | Mean<br><br>(SD) | n = 58<br><br>1.3 (3.8) | n = 63<br><br>14.2 (20.1) | MD –12.90, 95% CI –17.96 to –7.84 |
| Kim 2011b | **Urine leakage score**<br><br>This is calculated based on the self-reported 1-week urinary diary (score of 0–4; with 0 = no urine leakage, 1 = less than once per week, 2 = once per week, 3 = 2 or 3 times per week, and 4 = every day). No information was found on the psychometric properties of this instrument. | Mean score (SD) | n = 59<br><br>3.0 (2.0) | n = 61<br><br>4.4 (1.6) | MD –1.4, 95% CI (–2.1 to –0.8) |
| Sran 2016 | **Number of urinary leakage episodes in 24 hours (at 1 year)** | Mean<br><br>(SD) | n = 24<br><br>0.49 (0.57) | n = 24<br><br>2.15 (2.87) | MD –1.66, 95% CI –2.83 to –0.49 |
| Yoon 2003 | **Urinary incontinence score**<br><br>Sum of scores from 5-point Likert scales regarding severity of leakage with 18 prespecified activities associated | Mean (SD) | n = 13<br><br>10.8 (6.2) | n = 12<br><br>14.2 (3.6) | MD –3.4, 95%<br><br>CI –7.6 to 0.8 |

(Continued)

| | |
|---|---|
| | with urine loss. No information was found on the psychometric properties of this instrument. |

CI: confidence interval; MD: mean difference; n: number of participants; PFMT: pelvic floor muscle training; SD: standard deviation.

## Appendix 6. Other non-specific quality of life outcomes

| Study ID | Outcome | Measure | Subscale | PFMT | Control | Difference |
|---|---|---|---|---|---|---|
| Burgio 1998 | **Hopkins Symptom Checklist for psychological distress (SCL-90-R)**<br>A 90-item self-administered questionnaire with 9 clinical subscales aggregated into a total score: the Global Severity Index. A score of 50 is normal. A score of more than 63 is a 'case' on any of the subscales. | Mean score (SD) | All | n = 57 | n = 46 | |
| | | | Somatisation | 51.8 (11.4) | 49.8 (13.0) | MD 2.0, 95% CI −2.8 to 6.8 |
| | | | Obsessive/compulsive | 53.8 (13.9) | 55.4 (11.0) | MD −1.6, 95% CI −5.7 to 2.5 |
| | | | Interpersonal sensitivity | 49.5 (12.0) | 49.2 (11.3) | MD 0.3, 95% CI −4.3 to 4.9 |
| | | | Depression | 51.5 (11.5) | 51.4 (11.2) | MD 0.1, 95% CI −6.7 to 1.9 |
| | | | Anxiety | 46.1 (14.6) | 45.8 (12.9) | MD 0.3, 95% CI −6.7 to 1.9 |
| | | | Hostility | 44.9 (10.8) | 47.3 (11.2) | MD −2.4, 95% CI −6.7 to 1.9 |
| | | | Phobia | 47.1 (11.2) | 45.1 (8.5) | MD 2.0, 95% CI −2.0 to 6.0 |
| | | | Paranoia ideation | 45.8 (10.9) | 47.2 (12.0) | MD −1.4, 95% CI −5.9 to 3.1 |
| | | | Psychoticism | 49.2 (11.7) | 49.6 (10.3) | MD −0.4, 95% CI −4.8 to 4.0 |
| | | | Global severity | 50.8 (12.8) | 51.4 (10.9) | MD −0.6, 95% CI −4.7 to 3.5 |
| Bø 1999 | The Norwegian version of the Quality of Life Scale (QoLS-N)<br>A 16-item scale used in populations with chronic illness. Uses a 7-point satisfac- | Mean total score (SD) | NA | n = 25<br>90.1 (9.5) | n = 30<br>85.2 (12.1) | MD 4.9, 95% CI −1.1 to 10.9 |

(Continued)

| | | | | | | |
|---|---|---|---|---|---|---|
| | tion scale per item whereby a higher score indicates a higher quality of life. | | | | | |
| Kargar Jahromi 2013 | **Rosenberg's self-esteem evaluation**<br>The Rosenberg evaluation aims to establish the participant's feelings about 10 sentences or comments (4: strongly agree, 3: agree, 2: disagree and 1: strongly disagree). Total points are obtained by adding the points for all 10 questions (from 10 to 40 points), which indicates the minimum and maximum self-esteem. No details was given on its psychometrics. | Mean score (SD) | NA | n = 24<br>27.66 (4.00) | n = 24<br>22.38 (5.04) | MD 5.28, 95% CI 2.71 to 7.85 |
| Sran 2016 | Geriatric Self-Efficacy Index for Urinary Incontinence after treatment | Mean score (SD) | NA | n = 24<br>0.65 (0.23) | n = 24<br>0.46 (0.23) | MD 0.19, 95% CI 0.06 to 0.32 |
| | Geriatric Self-Efficacy Index for Urinary Incontinence at 1 year | Mean score (SD) | NA | n = 24<br>0.65 (0.80) | n = 24<br>0.50 (0.26) | MD 0.15, 95% CI 0.02 to 0.28 |

CI: confidence interval; MD: mean difference; n: number of participants; NA: not applicable; SD: standard deviation.

## Appendix 7. Incontinence aids

| Study ID | Outcome | Measure | PFMT | Control | Difference |
|---|---|---|---|---|---|
| Asklund 2017 | No use of pads in 4 weeks | Number (%) | 23 (37.7) | 14 (23.3) | RR 1.62, 95% CI 0.92 to 2.83 |
| | Less than once per week | | 13 (21.3) | 11 (18.3) | RR 1.16, 95% CI 0.57 to 2.39 |
| | 1–3 times per week | | 10 (16.4) | 13 (21.7) | RR 0.76, 95% CI 0.36 to 1.59 |
| | > 3 times per week but not daily | | 5 (8.2) | 4 (6.7) | RR 1.23, 95% CI 0.35 to 4.36 |
| | 1 aid per day | | 8 (13.1) | 12 (20) | RR 0.66, 95% CI 0.29 to 1.49 |
| | >1 per day | | 2 (3.3) | 6 (10) | RR 0.33, 95% CI 0.07 to 1.56 |

CI: confidence interval; PFMT: pelvic floor muscle training; RR: risk ratio.

## Appendix 8. Other pad or paper towel test

| Study ID | Outcome | Measure | PFMT | Control | Difference |
|---|---|---|---|---|---|

*(Continued)*

| | | | | | |
|---|---|---|---|---|---|
| Sran 2016 | 24-hour pad test at 1 year (g) | Mean (SD) | n = 24<br>2.8 (2.4) | n = 24<br>33.3 (63.6) | MD −30.50, 95% CI −55.97 to −5.03 |
| Aksac 2003 | 1-hour pad test (g) | Median (SD) | n = 20<br>2.1 (0.4) | n = 20<br>28.2 (3.7) | Not estimable |
| Bidmead 2002 | Short pad test, weight change from baseline (g) | Mean (SD) | n = 40<br>−9.6 (3.4) | n = 20<br>3.7 (1.2) | MD −13.3, 95% CI −23.1 to −3.4 |
| Diokno 2010 | Cough test (cm) | Mean (SD) | n = 23<br>12.6 (41.6) | n = 18<br>19.6 (48.8) | MD 25.30, 95% CI −2.9 to 53.5 |
| Miller 1998 | Paper towel test, wet area ($cm^2$) | Mean (SD) on medium cough | n = 13<br>0.4 (1.0) | n = 10<br>21.2 (44.8) | MD −20.8, 95% CI −46.5 to 4.9 |
| | | Mean (SD) on deep cough | n = 13<br>5.4 (15.3) | n = 10<br>26.8 (46.7) | MD −21.4, 95% CI (−50.0 to 7.2) |

CI: confidence interval; MD: mean difference; n: number of participants; PFMT: pelvic floor muscle training; SD: standard deviation.

## Appendix 9. Pelvic floor muscle training function assessment

| Assessments | PFMT outcomes and study ID | Outcome | Measure | PFMT | Control | Difference |
|---|---|---|---|---|---|---|
| **US measurements** | Carneiro 2010 | Transperineal US<br>Bladder neck mobility (mm) | Mean (SD) | n = 25<br>12.63 (4.35) | n = 25<br>17.53 (4.33) | MD −4.90, 95% CI −7.3 to −2.5 |
| | | Transperineal US<br>PFM thickness (mm) | Mean (SD) | n = 25<br>12.87 (1.02) | n = 25<br>10.74 (2.26) | MD 2.13, 95% CI 0.4 to 3.9 |
| | Celiker Tosun 2015 | Bladder base movement (mm) | Mean (SD) | n = 58<br>10.4 (6.5) | n = 63<br>5.4 (9.8) | MD 5.00, 95% CI 2.06 to 7.94 |
| | McLean 2013 | Vertical position of the bladder neck relative to the levator plate at rest (supine) (cm) | Mean (95% CI) | n = 15<br>2.04 (1.93 to 2.15) | n = 17<br>2.20 (2.10 to 2.30) | — |
| | | Vertical position of the bladder neck relative to the levator plate at rest (standing) (cm) | Mean (95% CI) | n = 15<br>0.99 (0.80 to 1.18) | n = 17<br>1.06 (0.72 to 1.40 | — |
| | | Displacement of the bladder neck during Valsalva (supine) (cm) | Mean (SD) | n = 15<br>2.02 (0.75) | n = 17<br>2.24 (1.06) | — |
| | | Displacement of the bladder neck during Valsalva (standing) (cm) | Mean (SD) | n = 15<br>1.60 (0.84) | n = 17<br>1.58 (0.64) | — |
| | | Displacement of the bladder neck during coughing (supine) (cm) | Mean (SD) | n = 15<br>0.96 (0.66) | n = 17<br>2.00 (0.76) | — |
| | | Displacement of the bladder neck during coughing (standing) (cm) | Mean (SD) | n = 15<br>0.88 (0.64) | n = 17<br>1.39 (0.60) | — |
| | | Urethral wall cross sectional area at 5 different points ($cm^{-2}$) | Mean (SD) | n = 15<br>1.21 (0.56) | n = 17<br>1.64 (0.35) | — |

*(Continued)*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | –1 | Mean (SD) | n = 15<br>1.21 (0.51) | n = 17<br>1.65 (0.35) | — |
| | | | Mid point | Mean (SD) | n = 15<br>1.26 (0.52) | n = 17<br>1.63 (0.31) | — |
| | | | 1 | Mean (SD) | n = 15<br>1.35 (0.56) | n = 17<br>1.67 (0.33) | — |
| | | | 2 | Mean (SD) | n = 15<br>1.40 (0.58) | n = 17<br>1.87 (0.43) | — |
| **Pressure measurements** | Aksac 2003 | Intravaginal<br>($cmH_2O$) | | Median (SD) | n = 20<br>37.5 (8.7) | n = 10<br>20.0 (3.9) | Non-estimable |
| | Beutten-muller 2010 | Intravaginal ($cmH_2O$) | | Mean (SD) | n = 25<br>Slow twitch<br>22.74 (5.65) | n = 25<br>Slow twitch 17.70 (5.86) | MD 5.04, 95% CI 1.9 to 8.2 |
| | | | | | Fast twitch<br>32.72 (10.34) | Fast twitch<br>28.09 (9.89) | MD 4.63, 95% CI –0.03 to 9.3 |
| | Bø 1999 | Intravaginal ($cmH_2O$) | | Mean (SD) | n = 25<br>19.2 (10.0) | n = 30<br>16.4 (9.8) | MD 2.8, 95%<br>CI –2.6 to 8.2 |
| | Celiker Tosun 2015 | Vaginal squeeze pressure<br>($cmH_2O$) | | Mean (SD) | n = 58<br>20.3 (15.5) | n = 63<br>16.2 (11.6) | MD 4.10, 95% CI –0.81 to 9.01 |
| | Firra 2013 | vaginal squeeze pressure<br>($cmH_2O$) | | Mean (SD)<br>SUI | n = 12<br>32.5 (18.5) | n = 9<br>26.1 (18.6) | MD 6.40, 95% CI –9.64 to 22.44 |

(Continued)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | UUI | n = 6<br>47.2 (22.7) | | n = 6<br>34.3 (25.5) | MD 12.90, 95% CI −14.42 to 40.22 |
| | Pereira 2011 | Intravaginal ($cmH_2O$) | Mean (SD) | Group PFMT<br>n = 15<br>37.13 (19.24) | Individual PFMT<br>n = 15<br>38.53 (19.34) | n = 15<br>11.91 (5.57) | MD 25.92, 95%<br>CI 18.45 to 33.0 |
| | Yoon 2003 | Mean pressure, intravaginal (mmHg) | Mean (SD) | n = 13<br>26.1 (12.5) | | n = 12<br>12.2 (5.3) | MD 13.9, 95%<br>CI 5.8 to 22.0 |
| | | Peak pressure, intravaginal (mmHg) | Mean (SD) | 39.7 (20.0) | | 19.9 (7.5) | MD 19.8, 95%<br>CI 7.1 to 32.5 |
| | | Duration of PFM contraction(s) | Mean (SD) | 14.5 (3.0) | | 5.9 (1.7) | MD 8.6, 95%<br>CI 6.6 to 10.6 |
| **Digital measurements** | Aksac 2003 | Intravaginal<br>Number of fingers not stated<br>Scale: 5-point scale | Median (SD) | n = 20<br>4.8 (0.4) | | n = 10<br>3.3 (0.6) | Not estimable |
| | Beutten-muller 2010 | Intravaginal<br>1 finger<br>Scale: Oxford | Mean (SD) | n = 25<br>Slow twitch<br>3.84 (0.8) | | n = 25<br>Slow twitch<br>2.95 (0.90) | MD 0.45, 95%<br>CI −0.02 to 0.92 |
| | | | | Fast twitch<br>3.80 (0.65) | | Fast twitch<br>2.86 (0.77) | MD 0.94, 95%<br>CI 0.6 to 1.3 |
| | Carneiro 2010 | Intravaginal<br>2 fingers<br>Scale: not stated | Mean (SD) | n = 25<br>3.20 (1.05) | | n = 25<br>2.50 (0.76) | MD 0.7, 95%<br>CI 0.2 to 1.21 |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| Castro 2008 | Intravaginal<br>Number of fingers not stated<br>Scale: Oxford | Mean (SD) | n = 26<br>3.6 (0.71) | n = 24<br>2.3 (1.07) | MD 1.30, 95% CI 0.79 to 1.81 |
| Celiker Tosun 2015 | P | Mean (SD) | n = 58<br>4.4 (0.9) | n = 63<br>2.6 (0.9) | MD 1.80, 95% CI 1.48 to 2.12 |
| | E | | n = 58<br>36.5 (17.7) | n = 63<br>25.9 (6.4) | MD 10.60, 95% CI 5.78 to 15.42 |
| | R | | n = 58<br>18.0 (6.2) | n = 63<br>7.6 (3.5) | MD 10.40, 95% CI 8.59 to 12.21 |
| | F | | n = 58<br>17.0 (5.8) | n = 63<br>8.9 (4.2) | MD 8.10, 95% CI 6.28 to 9.92 |
| Diokno 2010 | Intravaginal<br>Number of fingers not stated<br>Scale: not stated | — | n = 23 | n = 18 | — |
| | Pressure | Mean (SD) | 4.1 (1.1) | 3.8 (0.9) | MD 0.30, 95% CI −0.3 to 0.9 |
| | Displacement | Mean (SD) | 2.3 (1.3) | 2.1 (0.9) | MD 0.20, 95% CI −0.5 to 0.9 |
| | Duration | Mean (SD) | 7.1 (2.9) | 5.9 (3.1) | MD 1.2, 95% CI −0.7 to 3.1 |
| Miller 1998 | Intravaginal<br>Number of fingers not stated<br>Score: 0–21 | Mean (SD) | n = 13<br>10.4 (4.7) | n = 13<br>11.2 (5.1) | MD −1.1, 95% CI −5.1 to 2.9 |

(Continued)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Pereira 2011 | Intravaginal<br>2 fingers<br>Scale: 6-point modified Oxford scale | Mean (SD) | Group PFMT<br>n = 15<br>3.07 (0.70) | Individual PFMT<br>n = 15<br>2.73 (0.96) | n = 15<br>1.47 (0.52) | MD 1.43, 95% CI 1.0 to 1.46 |
| | Wells 1999 | Intravaginal<br>Number of fingers not stated<br>Scale: pressure and displacement digital score (4–12) | Mean | 8.8 | | 8.2 | Not estimable |
| **EMG measurements** | Bertotto 2017 | Initial EMG baseline | Mean (SD) | n = 15<br>16.3 (2.9) | | n = 14<br>13.78 (4.0) | MD 2.52, 95% CI −0.04 to 5.08 |
| | | Final EMG baseline | | n = 15<br>15.9 (2.4) | | n = 14<br>13.85 (3.7) | MD 2.05, 95% CI −0.24 to 4.34 |
| | | Duration of endurance contraction in seconds | | n = 15<br>6.8 (2.01) | | n = 14<br>2.35 (2.30) | MD 4.45, 95% CI 2.87 to 6.03 |
| | | Maximum voluntary contraction | | n = 15<br>20 (5.21) | | n = 14<br>15.9 (7.0) | MD 4.1, 95% CI −0.42 to 8.62 |
| | | Precontraction | | n = 15<br>0.67 (0.12) | | n = 14<br>0.21 (0.11) | MD 0.46, 95% CI 0.38 to 0.54 |
| | Burns 1993 | Intravaginal EMG<br>5 fast contractions | Mean (SD) | n = 38<br>3.0 (3.4) | | n = 40<br>3.5 (4.4) | MD −0.5, 95% CI −2.3 to 1.3 |
| | | Intravaginal EMG<br>5 sustained contractions | Mean (SD) | n = 33<br>1.8 (2.0) | | n = 34<br>2.0 (1.8) | MD −0.2, 95% CI −1.1 to 0.7 |
| | Carneiro 2010 | Intravaginal EMG | Mean (SD) | n = 25 | | n = 25 | MD 5.31, 95% |

(Continued)

| | | | | | |
|---|---|---|---|---|---|
| | 3 maximal contractions | | 13.56 (5.41) | 8.25 (5.70) | CI 2.23 to 8.39 |
| Wells 1999 | Intravaginal or intra-anal EMG<br>4 sustained and 4 short contractions | Mean | 48.8 | 24.2 | Not estimable |

CI: confidence interval; EMG: electromyography; MD: mean difference; n: number of participants; PFM: pelvic floor muscle; PFMT: pelvic floor muscle training; SD: standard deviation; US: ultrasound.

### Appendix 10. Number of women with interference with life due to urinary incontinence

| Study ID | Outcome | PFMT | Control | Difference |
|---|---|---|---|---|
| Bø 1999 | Number of women with interference with life due to UI | 17/25 | 25/30 | RR 0.82, 95% CI 0.60 to 1.12 |

CI: confidence interval; PFMT: pelvic floor muscle training; RR: risk ratio; UI: urinary incontinence.

## WHAT'S NEW

| Date | Event | Description |
|---|---|---|
| 3 October 2018 | New search has been performed | In this update, published in 2018, we made the following changes.<br>• Updated the search to February 2018 and added 10 new trials (Asklund 2017; Bertotto 2017; Celiker Tosun 2015; Ferreira 2014; Firra 2013; Kargar Jahromi 2013; Leong 2015; McLean 2013; Solberg 2016; Sran 2016).<br>• Completed a full risk of bias assessment for all included trials, including the addition of the 'blinding of participants and personnel' and 'selective reporting' domains.<br>• Assessed the quality of the body of evidence by adopting the GRADE approach.<br>• Added a Brief Economic Commentary to summarise the findings of eligible economic evaluations.<br>• Deleted 'treatment adherence' from the 'Summary of findings' tables because all review authors agreed adherence is a mediator rather than a pelvic floor muscle training outcome.<br>• Replaced the King's Health Questionnaire in the 'Summary of findings' table by a narrative report on all Grade A urinary incontinence-specific symptoms and quality of life (QoL) outcomes, as they give a better overview of combined urinary incontinence-specific symptoms and QoL outcomes.<br>• Added new Grade A symptoms and QoL questionnaires to the forest plots. |
| 3 October 2018 | New citation required but conclusions have not changed | The byline has been changed. |

## HISTORY

Protocol first published: Issue 1, 1999
Review first published: Issue 1, 2001

| Date | Event | Description |
|---|---|---|
| 13 May 2014 | New search has been performed | In this update, seven new trials have been added (Beuttenmuller 2010; Carneiro 2010; Diokno 2010; Kim 2011a; Kim 2011b; Pereira 2011; Sar 2009). One previously included trial has been removed because the control group was deemed to be receiving a form of active treatment (van Leeuwen 2004). Full risk of bias assessment has been completed for all trials. Data from 'Other data' tables have been incorporated into other sections. Quality of evidence was assessed by adopting the GRADE approach. |
| 13 May 2014 | New citation required and conclusions have changed | In this update, seven new trials have been added (Beuttenmuller 2010; Carneiro 2010; Diokno 2010; Kim 2011a; Kim 2011b; Pereira 2011; Sar 2009). One previously included trial has been removed because the control group was deemed to be receiving a form of active treatment (van Leeuwen 2004). Full risk of bias assessment has been completed for all trials. Data from 'Other data' tables have been incorporated into other sections. Quality of evidence was assessed by adopting the GRADE approach. |

## CONTRIBUTIONS OF AUTHORS

All three review authors were involved in all stages of the review. Chantale Dumoulin and Licia Cacciari wrote the first draft of this version of the updated review.

## DECLARATIONS OF INTEREST

CD: none known.
LPC: none known.
JHS: none known.

CD and JHS have both published trials investigating the effects of PFMT. Both trials were excluded from this review based on the participants (antenatal and postnatal women) or the comparison interventions (one type of PFMT versus another).

## SOURCES OF SUPPORT

### Internal sources

- University of Montreal, Canada.

### External sources

- National Institute for Health Research (NIHR), UK.

  This project was supported by the National Institute for Health Research (NIHR), via Cochrane Infrastructure, Cochrane Programme Grant or Cochrane Incentive funding to the Cochrane Incontinence Group. The views and opinions expressed therein are those of the review authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, National Health Service or the Department of Health. The NIHR is the largest single funder of the Cochrane Incontinence Group.
- Chantale Dumoulin was funded by the Canadian Research Chair of the Canadian Institute of Health Research, Canada.
- Licia P Cacciari, Postdoctoral Research Fellow, Centre de recherche de l'Institut de Gériatrie de Montréal, Canada.

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

For this update, published in 2018, we made the following changes.

- Risk of bias assessment has been conducted for random sequence generation, allocation concealment, blinding of participants and personnel, incomplete outcome data, selective reporting, blinding of outcome assessment and baseline comparability. Risk of bias figures and tables have been added in the last two review updates.
- GRADE assessment has been conducted for the two last review updates. Substantive changes have been made to provide more details about the methodology since the publication of the protocol, in line with current Cochrane standards.

- For this update, treatment adherence was removed from 'Summary of findings' tables as it is considered to be a mediator rather than an outcome of PFMT. For the next review update, we will remove treatment adherence from the outcome list and report it in the PFMT protocol (Appendix 3).
- King's Health Questionnaire was replaced in the 'Summary of findings' table by a narrative report on all Grade A UI-specific symptoms and quality of life outcomes, as they will give a better overview of combined UI-specific symptoms and quality of life outcomes.
- New Grade A symptoms and quality of life questionnaires have been added to the forest plots.
- Health economic outcomes have been removed as a review outcome and from the 'Summary of findings' tables. In their place, a Brief Economic Commentary has been added.

## INDEX TERMS

### Medical Subject Headings (MeSH)

*Pelvic Floor; Biofeedback, Psychology; Exercise Therapy [*methods]; Muscle Contraction [*physiology]; Perineum; Quality of Life; Randomized Controlled Trials as Topic; Urinary Incontinence [*rehabilitation]; Urinary Incontinence, Stress [rehabilitation]

### MeSH check words

Female; Humans