CHAPTER 4

TWO-YEAR EFFECTS AND COST-EFFECTIVENESS OF PELVIC FLOOR MUSCLE TRAINING IN MILD PELVIC ORGAN PROLAPSE: A RANDOMISED CONTROLLED TRIAL IN PRIMARY CARE

Chantal M.C.R. Panman, Marian Wiegersma, Boudewijn J. Kollen,
Marjolein Y. Berger, Yvonne Lisman-Van Leeuwen, Janny H. Dekker

British Journal of Obstetrics and Gynaecology, 2016
ABSTRACT

Objective
To compare effects and cost-effectiveness of PFMT and watchful waiting in women with prolapse.

Design, setting and population
A randomised controlled trial in Dutch general practice. Women (≥55 years) with symptomatic mild prolapse who were identified by screening were included.

Methods
Linear multilevel analysis.

Main outcome measures
Primary outcome was change of pelvic floor symptoms (PFDI-20) during 24 months. Secondary outcomes were condition-specific and general quality of life, costs, sexual functioning, prolapse stage, pelvic floor muscle function and women's perceived improvement of symptoms.

Results
PFMT (n = 145) resulted in a 12.2-point (95% CI 7.2 to 17.2, p < 0.001) greater improvement in PFDI-20 score during 24 months compared with watchful waiting (n = 142). Participants randomised to PFMT more often reported improved symptoms (43% versus 14% for watchful waiting). Direct medical costs per person were €330 for PFMT and €91 for watchful waiting but costs for absorbent pads were lower in the PFMT group (€40 versus €77). Other secondary outcomes did not differ between groups. Post-hoc subgroup analysis demonstrated that PFMT was more effective in women experiencing higher pelvic floor symptom distress at baseline.

Conclusion
PFMT resulted in greater pelvic floor symptom improvement compared with watchful waiting. The difference was statistically significant, but below the presumed level of clinical relevance (15 points). PFMT more often led to women's perceived improvement of symptoms, lower absorbent pads costs, and was more effective in women experiencing higher pelvic floor symptom distress. Therefore, PFMT could be advised in women with bothersome symptoms of mild prolapse.
INTRODUCTION

Pelvic organ prolapse is characterised by a descent of the vaginal wall or apex from its normal anatomical position. In a community survey, 75% of Dutch women aged 45–85 years had a prolapse on physical examination. Given the aging population and that prolapse is more common in the elderly, this prevalence is likely to increase. Women with prolapse present with a variety of pelvic floor symptoms, including vaginal bulging, a feeling of pelvic pressure or heaviness, pelvic pain, and urinary or faecal incontinence or obstruction. Prolapse can affect daily activities, sexual function, and quality of life. Treatment typically involves either conservative management or surgery, with the former preferred in cases of a mild prolapse or if the woman is not a good candidate for surgery because of co-morbidity. Conservative treatments comprise PFMT, which aims to improve pelvic floor muscle function, and mechanical devices (vaginal pessaries), which aim to redress the prolapse. There are two main hypotheses on how PFMT may be effective in the prevention and treatment of prolapse: (1) descent of the pelvic floor is prevented by teaching women to consciously contract their pelvic floor muscles before and during any increase in abdominal pressure and (2) ‘stiffness’ and structural support of the pelvic floor are built up by performing regular muscle strength training over time. Although studies have shown short-term symptom improvement when comparing PFMT with watchful waiting or lifestyle advice for stage 1 or 2 prolapse, evidence on the medium-term and long-term clinical effects and cost-effectiveness is lacking. The aim of this study was to compare the effects and cost-effectiveness of PFMT and watchful waiting in older women with a symptomatic mild prolapse in primary care during a 2-year follow up.

METHODS

Design and setting
This randomised controlled trial was conducted in 15 general practices in the Netherlands, with enrolment taking place between 14 October 2009 and 19 October 2012. The last date of follow-up was 28 November 2014. All participants gave written informed consent. The study protocol and short-term results have been reported previously.

Participants
The participating general practitioners selected all women aged ≥55 years registered in their practice who did not meet the exclusion criteria: current prolapse treatment
or treatment in the past year, current treatment for another (uro)gynaecological disorder, malignancy of pelvic organs, impaired mobility, severe or terminal illness, cognitive impairment, or insufficient command of the Dutch language. Eligible women received a postal five-item screening questionnaire on pelvic floor symptoms (urinary incontinence, vaginal bulging, pelvic heaviness/pressure and vaginal splinting to start or complete micturition or defecation). Women who screened positive for one or more pelvic floor symptoms were invited to a baseline assessment, which was conducted by the research physician. This assessment comprised a standardised interview, completion of the study questionnaires, and a physical examination. The examination included an assessment of the degree of prolapse using the POP-Q system.13 Women in whom the leading edge of the prolapse stayed above the hymeneal remnants (POP-Q stage 1 or mild stage 2) were considered to have a mild prolapse and were eligible for inclusion and randomisation.

Randomisation and blinding
The research physician randomly allocated women to PFMT or watchful waiting (1:1 ratio), using an external computer system (accessible by telephone). Block randomisation with variable block sizes was used. The research physician enrolling participants was blinded to both allocation sequence and ordering of the blocks and their sizes. The participants, physiotherapists and physicians were not blinded to group allocation. Research physicians performing the follow-up assessments were blinded to all responses on the questionnaires, as well as to the outcomes of the previous physical examination. The analyses were performed by a researcher who was blinded to group allocation.

Interventions

Pelvic floor muscle training
All participating physiotherapists had to be registered with the Dutch Pelvic Physiotherapists’ Organisation to improve the level of standardisation between interventions. Therapy was introduced with an explanation and description of the pelvic floor, and participants were instructed how to contract and relax their pelvic floor muscles. Participants who were not able to contract or relax their pelvic floor muscles were first taught how to do this by giving feedback during digital palpation. In case of insufficient control of pelvic floor muscles after feedback during digital palpation the pelvic physiotherapist could use myofeedback or electrical stimulation. Everyone then started with the same exercise programme, which was subsequently tailored to individual needs (Appendix I). All participants were taught the correct technique.
for contracting the pelvic floor muscles before and during increases in abdominal pressure (‘the Knack’), and received information about toilet habits and lifestyle (diet, smoking, body weight). In participants with overactive pelvic floor muscles, emphasis was placed on relaxation instead of contraction, using both general relaxation exercises and specific pelvic floor relaxation exercises. Participants had face-to-face contact with their physiotherapist and were encouraged to continue practicing at home three to five times a week for two to three times on each day. Treatment modalities and the number of sessions were provided by the pelvic physiotherapist and were recorded for each participant. Women were allowed to consult their physician for any symptoms of prolapse during the study period. These consultations and any resulting treatments were recorded at follow up.

**Watchful waiting**

Participants randomised to watchful waiting did not receive active treatment in the context of the trial. Before randomisation, women received information on pelvic anatomy and pelvic floor muscle function by using illustrated leaflets. In addition, they were informed about the degree of their prolapse and the function of their pelvic floor muscles. Women were allowed to consult their physician for prolapse-related symptoms during the study period, and any details were recorded at follow up.

**Follow up**

Follow up was by questionnaire and a physical examination at 3, 12 and 24 months after the start of the intervention. Participants who were not able to attend the follow-up appointment were asked to complete questionnaires sent by post. In addition, non-attenders were phoned to check if they had received any consultations or treatments for prolapse symptoms and to enquire about their use of absorbent pads.

**Outcome measures**

The primary outcome was distress of pelvic floor symptoms measured with the PFDI-20.14 This is a validated patient-completed questionnaire with 20 questions about prolapse, anorectal and urinary symptoms. Total PFDI-20 score ranges from 0 to 300, with higher scores indicating higher symptom burden.

Secondary outcomes were the three subscales of the PFDI-20, assessing distress of prolapse symptoms with the POPDI-6, colorectal/anal symptoms with the CRADI-8, and urinary symptoms with the UDI-6, respectively.14 In addition, condition-specific and general quality of life were measured by the PFIQ-714 and by physical (PCS-12) and mental (MCS-12) component scores of the MOS-SF-12.15 Sexual functioning was assessed using the PISQ-12.16 At 24 months, a five-point scale was used to assess
women’s perceived Global Perception of Improvement (GPI) from the start of the study (much better, better, the same, worse, or much worse). Details of consultations, prolapse-related treatments, and the use of absorbent pads were recorded during each follow-up appointment and used for the cost-effectiveness analysis. The three-level version of the EuroQol health status measure (EuroQol 5D-3L) was used to collect data for the cost-utility evaluation. The degree of prolapse was assessed with the POP-Q, using the ordinal scale for stages of prolapse (0–4) for each compartment (anterior, posterior, apical). In the POP-Q system the degree of prolapse of the anterior vaginal wall (point Ba), the posterior vaginal wall (point Bp) and the uterus or vaginal vault (when there is a history of hysterectomy) (point C) is measured in centimetres using the hymenal remnants as a reference point. The overall POP-Q stage is equal to the POP-Q stage of the most severely prolapsed compartment. The pelvic floor muscle function was assessed by digital vaginal palpation in the supine position and was categorised as normal, underactive, overactive or inactive according to the International Continence Society classification. Research physicians were trained to perform a POP-Q measurement and evaluation of the pelvic floor muscles by an experienced urogynaecologist.

**Statistical analyses**

Based on an estimated clinically relevant difference in change of PFDI-20 score of 15 points between groups, with a standard deviation of 36 points, type I error of 5%, power of 80%, and dropout rate of 15% after 2 years, we needed to enrol 216 women. All statistical analyses were performed two-sided at a 5% significance level.

**Primary and secondary outcomes**

In this study outcome variables were repeatedly measured over time and compared between groups. The analysis of such data requires a statistical approach that corrects for the correlated information generated by repeated measurements. We conducted a multi-level analysis (MLA) using MLWIN 2.29 (Centre for Multilevel Modelling, University of Bristol, UK). Multilevel models are hierarchical systems that estimate regression coefficients and their variance components while at the same time correcting for the dependency of information. The first level was defined as time of measurement, the second level as patient. The restricted iterative generalised least squares algorithm was used to estimate the regression coefficients, and the Wald test was used to obtain a p-value for each regression coefficient. A linear MLA model with a fixed and random intercept was used to test whether the treatment groups differed longitudinally with respect to the primary outcome variable PFDI-20 and secondary outcome variables POPDI-6, CRADI-8, UDI-6, PFIQ-7, PCS-12, MCS-12.
and PISQ-12. For each equation two intention-to-treat analyses were conducted, the difference being the adjustment of the questionnaire and POP-Q (stage 1 or stage 2) baseline scores. The linear MLA for the PFDI-20 was repeated in two post-hoc subgroup analyses of women with the top 25% and 50% of PFDI-20 baseline scores. In addition, supplementary analyses were performed to test the relationship between the degree of prolapse of the anterior vaginal wall (point Ba), the posterior vaginal wall (point Bp) and the uterus or vaginal vault (point C) between treatment groups, adjusted for its baseline measurements. We did not impute missing data as replacing missing values in longitudinal data sets is considered redundant. The assumptions of normality and homogeneity of the variance were assessed by inspecting normal probability plots and plots of standardised residuals versus predicted values. In the event of noncompliance, a square or square root transformation of one of the baseline variables was implemented.

Chi-square tests were used to determine between-group differences in the GPI, overall change in POP-Q stage, and change of pelvic floor muscle function between baseline and 24 months. For these statistical analyses, IBM SPSS for Windows, version 23.0 (IBM Corp., Armonk, NY, USA) was used.

Cost-effectiveness

Direct medical costs were the costs for physical therapy, consultations with general practitioners and medical specialists, absorbent pads, and other costs, such as other prolapse-related treatments, like pessaries, PFMT and prolapse surgery. Costs were valued according to Dutch guidelines at the 2013 price level. In a cost-effectiveness analysis the balance between the costs and the effects of PFMT was evaluated compared with watchful waiting. In this cost-effectiveness analysis the incremental costs per PFDI-20 point gained, were assessed and expressed as an Incremental Cost-Effectiveness Ratio (ICER). In the ICER, the difference in effect, based on the primary outcome, was divided by the difference in costs between both interventions. In addition, a cost–utility analysis (represented by quality-adjusted life years [QALYs]) was performed to evaluate the balance between costs and QALYs, using the EQ-5D defined utility scores, based on the UK tariff. For reasons of comparability, e.g. the fact that the UK tariff is more frequently mentioned in the literature, compared to the Dutch tariff, we based our utility scores on the UK tariff. The outcome of the cost–utility analysis is presented as an Incremental Cost–Utility Ratio (ICUR). Bootstrap analysis (5000 replications of the trial data) was performed to estimate an alternative confidence interval for the cost difference and the point estimate of the ICER. Cost-effectiveness planes were constructed to visualise the uncertainty surrounding the ICER and ICUR.
RESULTS

Figure 1 shows the participant flow in the study. In total, 287 women were randomised to PFMT (n = 145) or watchful waiting (n = 142), and their baseline characteristics are shown in Table 1. For participants who completed PFMT, the median time of treatment was 16 weeks (IQR 12–24) and the median number of treatments was 7 (IQR 5–9). Myofeedback was used in 23 participants (16%) and electrical stimulation in 11 (8%). No adverse effects were reported for PFMT.

Primary and secondary outcomes

The questionnaire scores are shown in Table 2. Participants in the PFMT group showed a 12.2-point (95% CI 7.2 to 17.2, p < 0.001) greater improvement of pelvic floor symptoms (PFDI-20 score) during 24 months compared with participants in the watchful waiting group (Table 3).

Analyses of the PFDI-20 subscales showed a significant difference between groups in prolapse (POPDI-6), anorectal (CRADI-8) and urinary (UDI-6) symptoms in favour of the PFMT group (Table 3). No differences in change were observed between the groups in either sexual functioning (PISQ-12) or the condition-specific (PFIQ-7) and general quality of life (MCS-12 and PCS-12) during 24 months (Table 3).

After 24 months, more women in the PFMT group reported an overall improvement of symptoms (43% versus 14% for watchful waiting). In contrast, in the watchful waiting group more women reported that their symptoms remained unchanged (70% versus 52% for PFMT) or got worse (15% versus 5% for PFMT) (Table S1).

There was no difference between groups in the change of the degree of prolapse of the anterior vaginal wall (point Ba) (β = 0.1, 95% CI -0.1 to 0.3, p = 0.22), the uterus (point C) (β = −0.04, 95% CI -0.3 to 0.2, p = 0.75), and the posterior vaginal wall (point Bp) (β = 0.02, 95% CI -0.1 to 0.1, p = 0.77). There was no difference between groups in the proportion of women that showed either improvement or deterioration of ≥1 POP-Q stage, or in whom the POP-Q stage remained the same between baseline and 24 months (Table S2). The change in pelvic floor muscle function between baseline and 24 months was comparable in both groups (Table S3).

In the subgroup of women with the top 50% PFDI-20 baseline score (n = 139, mean PFDI-20 baseline score 90.6, SD 28.1), PFMT led to a 13.8-point (15%) (95% CI 5.7 to 21.9, p = 0.001) greater improvement of pelvic floor symptoms (PFDI-20 score) during 24 months compared with watchful waiting. For women with the top 25% PFDI-20 baseline score (n = 72, mean PFDI-20 baseline score 109.9, SD 26.3) this was 27.5 points (25%) (95% CI 14.6 to 40.4, p < 0.001) in favour of participants in the PFMT group.
Urinary incontinence, vaginal bulging, pelvic pressure/heaviness and vaginal splinting to start or complete micturition or defection. † In addition, three women received pessary treatment for prolapse from their own physician.

**FIGURE 1** FLOWCHART OF PARTICIPANTS THROUGH STUDY

* Women assessed for eligibility
  - 5411 Women assessed for eligibility
    - 946 Excluded (by primary care physician)
    - 946 Exclusion criteria
  - 4465 Screening questionnaire*  
    - 3751 Excluded
    - 1684 Not meeting inclusion criteria
    - 378 Declined to participate
    - 1689 No response
  - 714 Baseline assessment
    - 179 No prolapse
    - 170 Advanced prolapse
  - 365 Mild prolapse
    - 78 Declined to participate
  - 287 Randomized

**Randomized to PFMT**
- 145 Randomized to PFMT
- 134 Received PFMT as randomized†
- 11 Did not receive PFMT as randomized
  - 2 Personal circumstances
  - 4 Long waiting list at pelvic physiotherapist
  - 5 Unknown reasons

**Randomized to watchful waiting**
- 142 Randomized to watchful waiting
- 142 Received watchful waiting as randomized
- 0 Did not receive watchful waiting as randomized

**Lost to follow-up**
- 15 Lost to follow-up
  - 1 Health problems
  - 1 Personal circumstances
  - 2 Moved away
  - 4 Long waiting list at pelvic physiotherapist
  - 7 Unknown reasons
  - 27 Discontinued PFMT
  - 12 Personal circumstances
  - 6 Lack of motivation for PFMT
  - 1 Disappointing results of PFMT
  - 2 Found PFMT too time consuming
  - 1 Moved away
  - 1 Other prolapse treatment
  - 4 Unknown reasons

**Lost to follow-up**
- 7 Lost to follow-up
  - 2 Health problems
  - 1 Personal circumstances
  - 1 Moved away
  - 1 Deceased
  - 2 Unknown reasons
  - 10 Discontinued watchful waiting
  - 6 PFMT
  - 3 Pessary treatment
  - 1 Surgery (posterior wall repair)

**Included in questionnaire analysis**
- 145 at baseline
- 118 at 3 months
- 117 at 12 months
- 130 at 24 months

**Included in physical examination analysis**
- 145 at baseline
- 105 at 3 months
- 102 at 12 months
- 103 at 24 months
Cost-effectiveness

Direct medical costs over the 2-year study, per person, amounted to €330 in the PFMT group and €91 in the watchful waiting group. The mean difference was therefore €239 per person (95% CI €161 to €319). This difference can be explained by the difference in costs for physical therapy (€254 for PFMT versus €0 for watchful waiting), visits to the general practitioner and medical specialist (€24 for PFMT versus €11 for watchful waiting), and absorbent pad use (€40 for PFMT versus €77 for watchful waiting).
### TABLE 2 BASELINE AND FOLLOW-UP SCORES OF QUESTIONNAIRES

<table>
<thead>
<tr>
<th></th>
<th>PFMT</th>
<th>Watchful waiting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>3 months</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>PFDI-20</td>
<td>140</td>
<td>65.2 ± 39.9</td>
</tr>
<tr>
<td>POPDI-6</td>
<td>145</td>
<td>15.5 ± 13.4</td>
</tr>
<tr>
<td>CRADI-8</td>
<td>144</td>
<td>17.2 ± 15.3</td>
</tr>
<tr>
<td>UDI-6</td>
<td>141</td>
<td>32.4 ± 19.7</td>
</tr>
<tr>
<td>PFIQ-7</td>
<td>139</td>
<td>22.1 ± 39.9</td>
</tr>
<tr>
<td>PISQ-12*</td>
<td>64</td>
<td>35.5 ± 5.3</td>
</tr>
<tr>
<td>PCS-12</td>
<td>122</td>
<td>45.1 ± 10.5</td>
</tr>
<tr>
<td>MCS-12</td>
<td>122</td>
<td>52.7 ± 8.5</td>
</tr>
</tbody>
</table>

*PISQ-12 scores only available for women who were sexually active
### TABLE 3 DIFFERENCES IN QUESTIONNAIRES BETWEEN PFMT AND WATCHFUL WAITING DURING 24 MONTHS

<table>
<thead>
<tr>
<th></th>
<th>Mean (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFDI-20</td>
<td>7.4 (-0.5 to 15.3)</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>12.2 (7.2 to 17.2)</td>
</tr>
<tr>
<td>POPDI-6</td>
<td>1.9 (-0.7 to 4.5)</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>2.9 (2.7 to 3.1)</td>
</tr>
<tr>
<td>CRADI-8</td>
<td>1.8 (-1.3 to 4.9)</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>2.2 (0.3 to 4.2)</td>
</tr>
<tr>
<td>UDI-6</td>
<td>3.7 (-0.1 to 7.6)</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>5.7 (3.2 to 8.2)</td>
</tr>
<tr>
<td>PFIQ-7</td>
<td>-2.7 (-8.9 to 3.5)</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>-0.9 (-5.4 to 3.5)</td>
</tr>
<tr>
<td>PISQ-12\§</td>
<td>0.4 (-1.4 to 2.3)</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>-0.2 (-1.2 to 0.8)</td>
</tr>
<tr>
<td>PCS-12</td>
<td>0.8 (-1.4 to 3.1)</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>-0.2 (-1.8 to 1.3)</td>
</tr>
<tr>
<td>MCS-12</td>
<td>0.3 (-1.4 to 1.9)</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Adjusted*</td>
<td>0.1 (-1.2 to 1.5)</td>
</tr>
</tbody>
</table>

\*Adjusted for baseline questionnaire score and baseline POP-Q stage (1 or 2); †Square transformation of PFDI-20 baseline score; ‡Square root transformation of PFIQ-7 baseline score; §PISQ-12 scores only available for women who were sexually active.

The ICER of the PFDI-20 change score was €43 (95% CI €18 to €237), meaning that to gain one additional point on the PFDI-20 in the PFMT group (compared with watchful waiting) €43 had to be invested. Most of the 5000 replications of the bootstrap simulation (i.e. 98% of them) were in the northeast quadrant of the cost-effectiveness plane, indicating that they represented a better outcome and higher costs for the PFMT group (Figure S1).

With regard to the utility scores, both groups lost QALYs (0.067 in the watchful waiting group and 0.061 in the PFMT group). This resulted in an ICUR of €31,983 (95% CI -€76,652 to €88,078), meaning that an additional €31,983 had to be invested in the PFMT group for each QALY gained over the watchful waiting group. Of the 5000 bootstrap replications 55% was located in the northeast quadrant of the cost-effectiveness plane, indicating more effect for more money. The remaining 45% was located in the northwest quadrant, indicating less effect for more money (Figure S2).
**TABLE S1** SELF-REPORTED CHANGE OF SYMPTOMS BETWEEN BASELINE AND 24 MONTHS

<table>
<thead>
<tr>
<th></th>
<th>PFMT n/N (%)</th>
<th>Watchful waiting n/N (%)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Much better</td>
<td>10/129 (8)</td>
<td>3/130 (2)</td>
<td></td>
</tr>
<tr>
<td>Better</td>
<td>45/129 (35)</td>
<td>16/130 (12)</td>
<td></td>
</tr>
<tr>
<td>The same</td>
<td>67/129 (52)</td>
<td>91/130 (70)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Worse</td>
<td>7/129 (5)</td>
<td>20/130 (15)</td>
<td></td>
</tr>
<tr>
<td>Much worse</td>
<td>0/129 (0)</td>
<td>0/130 (0)</td>
<td></td>
</tr>
</tbody>
</table>

*p-value* indicates significance of difference between groups.

**TABLE S2** CHANGE OF POP-Q STAGE BETWEEN BASELINE AND 24 MONTHS

<table>
<thead>
<tr>
<th></th>
<th>PFMT n/N (%)</th>
<th>Watchful waiting n/N (%)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement (≥1 POP-Q stage)</td>
<td>28/101 (28)</td>
<td>20/116 (17)</td>
<td></td>
</tr>
<tr>
<td>The same</td>
<td>58/101 (57)</td>
<td>72/116 (62)</td>
<td>0.14</td>
</tr>
<tr>
<td>Deterioration (≥1 POP-Q stage)</td>
<td>15/101 (15)</td>
<td>24/116 (21)</td>
<td></td>
</tr>
</tbody>
</table>

*p-value* indicates significance of difference between groups.

**DISCUSSION**

**Main findings**

In women aged 55 years and over who were screened for symptomatic mild prolapse, PFMT resulted in a 12.2-point greater reduction of pelvic floor symptoms than watchful waiting. This improvement was associated with extra costs of €239 per patient in the PFMT group. The observed difference between groups was statistically significant but below the a priori presumed minimal clinically important difference of 15 points. We found no differences between groups in change in sexual functioning, condition-specific and general quality of life, nor in changes related to function of the pelvic floor muscles, and degree of prolapse.

**Strengths and limitations of this study**

The strengths of this study are its long follow up, its focus on patient-reported outcomes, and the low drop-out rate. Despite these strengths, a number of limitations deserve attention.

First, we screened and enrolled participants by general practice. This meant that
FIGURE S1  INCREMENTAL COST-EFFECTIVENESS PLANE FOR PFDI-20
5000 bootstrap replications for the mean difference between costs and PFDI-20

FIGURE S2  INCREMENTAL COST-EFFECTIVENESS PLANE FOR QALYS
5000 bootstrap replications for the mean difference between costs and QALYS
women aged ≥55 years from one practice were asked to participate in the study. For ethical reasons, once the sample size was reached, we could not refuse to include women we had already screened. We therefore included 287 women instead of the 216 estimated, this means that the study was overpowered and this might have increased the chance of finding a statistically significant difference.

Second, participants could not be blinded to treatment because of the nature of the intervention, so questionnaire responses may have been biased by their awareness of treatment allocation.

Third, although all physiotherapists were registered with the Dutch Pelvic Physiotherapists’ Organisation and participants were therefore trained according to the same basic exercise guidelines, programmes were tailored to each woman’s specific needs. Consequently, we cannot rule out differences in interventions and practices between physiotherapists. However, this reflects daily practice and should not invalidate our results.

Last, we chose to screen for prolapse because a considerable proportion of women with symptomatic prolapse do not consult their physician. Our aim was to screen sensitive to capture as many women with symptomatic prolapse as possible. We assumed that the women we recruited by screening would be comparable to women seeking help for prolapse as the main reasons for not seeking treatment are not related to symptom severity. Furthermore, the included women agreed to participate in a trial with an active treatment in one arm. Nevertheless, some participants had mild

<table>
<thead>
<tr>
<th></th>
<th>PFMT n/N (%)</th>
<th>Watchful waiting n/N (%)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal&gt;Normal</td>
<td>28/101 (28)</td>
<td>25/116 (22)</td>
<td></td>
</tr>
<tr>
<td>Underactive&gt;Normal</td>
<td>5/101 (5)</td>
<td>2/116 (2)</td>
<td></td>
</tr>
<tr>
<td>Overactive&gt;Normal</td>
<td>62/101 (61)</td>
<td>70/116 (60)</td>
<td></td>
</tr>
<tr>
<td>Normal&gt;abnormal</td>
<td></td>
<td></td>
<td>0.062</td>
</tr>
<tr>
<td>Normal&gt;Underactive</td>
<td>0/101 (0)</td>
<td>8/116 (7)</td>
<td></td>
</tr>
<tr>
<td>Normal&gt;Overactive</td>
<td>2/101 (2)</td>
<td>2/116 (2)</td>
<td></td>
</tr>
<tr>
<td>Abnormal&gt;Abnormal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underactive&gt;Overactive</td>
<td>4/101 (4)</td>
<td>9/116 (8)</td>
<td></td>
</tr>
<tr>
<td>Overactive&gt;Underactive</td>
<td>0/101 (0)</td>
<td>5/116 (4)</td>
<td></td>
</tr>
</tbody>
</table>

*Chi-square test
symptoms and, as a consequence, less potential for improvement. This might explain why the difference in change between PFMT and watchful waiting is smaller than expected. This is supported by the results of the post-hoc subgroup analysis, which showed that PFMT was more effective in women who experienced more distress of pelvic floor symptoms at the start of the study.

**Interpretation of results**

To date, no research has been published on the medium-term or long-term effects of PFMT on pelvic floor symptoms. Four trials compared PFMT with no intervention in women with stage 1 and stage 2 prolapse and showed favourable outcomes for PFMT on pelvic floor symptoms in the short term.\(^8\)\(^{-10}\)\(^{12}\) Two of these were pilot studies\(^8\)\(^,\)\(^10\) and one had serious methodological flaws.\(^9\) The fourth evaluated PFMT in the same population as included in this trial and found a statistically significant, but doubtfully clinically relevant, difference between PFMT and no treatment after 3 months.\(^12\)

Other studies, which included women with more severe degrees of prolapse (at or beyond the hymen) or without symptoms, also found a favourable effect of PFMT.\(^27\)\(^{-32}\)

Our study showed that PFMT results in a significantly greater improvement of pelvic floor symptoms compared with watchful waiting in a population that has been screened for symptomatic mild prolapse. The difference in change between groups was below the estimated level of clinical relevance (15 points). But, even if the difference between PFMT and watchful waiting is not clinically relevant at group level, there are several arguments to assume that PFMT is effective at least in some women with symptomatic mild prolapse. First, more women in the PFMT group reported improved symptoms than did in the watchful waiting group, suggesting that there was higher subjective, ‘patient relevant’ improvement. In the PFMT group, 43% of women reported that they were better, leaving 57% of women who said they were the same or worse, suggesting subgroups of responders and nonresponders. This finding is comparable to other studies investigating the effects of PFMT.\(^12\)\(^,\)\(^29\)

Second, costs for absorbent pads were lower for PFMT, suggesting a treatment benefit for those women. Last, women with more distress of pelvic floor symptoms at the start of the study seem to have more benefit from PFMT.

In agreement with the results of previous longitudinal studies on the natural history of prolapse, we found that some women had remission of prolapse while others had progression.\(^33\)\(^,\)\(^34\) In the majority of participants, the degree of prolapse did not change, suggesting that PFMT does not influence the natural course of prolapse during 24 months. A small difference was found with regard to QALYs. A generic quality-of-life scale, such as the EQ-5D, might not be sensitive enough to detect changes in quality of life in a condition like prolapse. No difference in life expectancy was expected between
groups, making it less useful to express the cost–utility in terms of QALYs. In view of the reasonable additional cost of PFMT (compared with watchful waiting), the decision whether or not to recommend PFMT for the treatment of mild prolapse should be based on clinical relevance and patient preference.

Many women experience difficulty in discussing prolapse with their physician because of a lack of knowledge about prolapse, its symptoms and treatment options or because of embarrassment due to symptoms. Our study suggests that at least some of the women with symptomatic mild prolapse who do not consult their physician on their own initiative can benefit from PFMT. Our advice to physicians would therefore be to adopt a proactive attitude regarding pelvic floor symptoms and prolapse. For example, they could ask briefly about distress of symptoms related to prolapse, micturition and defecation whenever a mild prolapse is found by coincidence (e.g. when performing a cervical smear). When there are only mild symptoms, watchful waiting can be advised after giving information about prolapse, whereas PFMT can be considered in case of more distress of pelvic floor symptoms.

CONCLUSIONS

This is the first study to investigate the 2-year effects of PFMT in a population of women aged 55 years and over with symptomatic mild prolapse, who were identified by screening a primary care population. PFMT resulted in a significantly larger reduction of pelvic floor symptoms compared with watchful waiting, but the difference between groups was below the estimated level of clinical relevance. Although further research is needed there are arguments to assume that PFMT is effective in at least some women with mild prolapse. Watchful waiting might be a good option when symptoms are mild, whereas PFMT might be considered in case of more severe distress. Future research should reveal the effects of PFMT in a population of women with bothersome symptoms and a wish for treatment. Furthermore, it would be useful to identify the characteristics of women who benefit from PFMT.
REFERENCES


14. Barber MD, Walters MD, Bump RC. Short forms of two conditionspecific quality-of-life


28. Kashyap R, Jain V, Singh A. Comparative effect of 2 packages of pelvic floor muscle training on


APPENDIX I PFMT LIFESTYLE AND TOILET ADVICE AND EXERCISE PROGRAM

Lifestyle advices
- Avoid heavy lifting.
- Avoid constipation: drink at least 1.5-2 liters of fluids daily and eat plenty of fiber (whole meal products, vegetables, fruit). Ask your general practitioner for a laxative if this is not sufficient.
- Use the Knack exercise before and during any increases abdominal pressure.
- If smoking: try to stop.
- If overweight: try to lose weight.

Toilet instructions
- Keep your feet flat on the floor.
- Micturition: sit up straight (concave back), relax your pelvic floor, do not strain.
- Defecation: go to the toilet if you feel the urge to defecate, sit down with a convex back, relax your pelvic floor and take your time. If necessary, strain slightly but keep breathing.

Basic exercises program
- Perform 3 series of 8-12 fast contractions of 1 second followed by 1 second of relaxation.
  Take 30-60 seconds of rest between series.
- Perform 3 series of 8-12 contractions of 6-10 seconds followed by 6-10 seconds of relaxation.
  Take 1-2 minutes of rest between series.
- Perform a (nearly) maximal contraction and hold this for 6-10 seconds, try to make 3-5 fast contractions of 1 second on top of the maximal contraction. Repeat this exercise 3 times and take 1-2 minutes of rest between exercises.¹

Examples of additional exercises
- Perform a step wise contraction: start at 30%, followed by 60%, up to 100% and also perform a stepwise relaxation: from 100% to 60% to 30% to complete relaxation. Repeat this exercise 3 times and take 1-2 minutes of rest between exercises.
- To focus on the ventral part of the pelvic floor: perform your basic exercises while sitting on a chair with a concave back and your feet positioned on the floor with your toes pointing inwards.
- To focus on the dorsal part of the pelvic floor: perform your basic exercises while sitting on a chair with a convex back and your feet positioned on the floor with your toes pointing outwards.
- Perform a contraction starting at the dorsal part of the pelvic floor, moving to the ventral part of the pelvic floor (like ‘closing a zipper’). Repeat this exercise 3 times and take 1-2 minutes of rest between exercises.

¹Not in participants with an overactive pelvic floor