



Unsupervised behavioral and pelvic floor muscle training programs for storage lower urinary tract symptoms in women: a systematic review

Chen Wu¹ · Diane K. Newman² · Mary H. Palmer¹

Received: 19 May 2020 / Accepted: 13 August 2020 / Published online: 1 September 2020
© The International Urogynecological Association 2020

Abstract

Introduction and hypothesis This study synthesized the effects of unsupervised behavioral and pelvic floor muscle training (B-PFMT) programs on outcomes relevant to women’s storage lower urinary tract symptoms (LUTS) and pelvic floor muscle strength (PFMS).

Methods PubMed, CINAHL, Web of Science, and PsycINFO were searched since their inception to August 6, 2019. Randomized controlled trials (RCTs) and quasi-experimental articles that enrolled community women aged 18 years and older and reported storage LUTS outcomes including symptoms, severity, impact, self-reported symptom improvement, and PFMS were screened and extracted. Risk of bias was evaluated, and a narrative synthesis approach was used to synthesize evidence.

Results Ten RCTs and three pretest-posttest articles were retrieved. Half of the RCTs had some concerns of bias; the remaining RCTs had high risk of bias. Three pretest-posttest articles had at least some risk of bias. Ten articles aimed to treat urinary incontinence (UI) primarily in middle-aged women, two aimed to prevent UI, and one aimed to improve PFMS in young continent women. Two months post-intervention, UI treatment effectiveness was observed, and cumulative effectiveness was evident in: (1) reduction in the number of incontinent episodes, (2) reduction in the International Consultation on Incontinence Questionnaire-Urinary Incontinence-Short Form and International Consultation on Incontinence Questionnaire-Lower Urinary Tract Symptoms Quality of Life scores, (3) increased patients’ global impression of improvement, and (4) improvement of PFMS.

Conclusions Unsupervised B-PFMT programs improve outcomes relevant to UI and PFMS in midlife community women who have UI. Their effects on UI prevention and other storage LUTS remain unclear.

Keywords Lower urinary tract symptoms · Urinary incontinence · Women · Pelvic floor · Exercise

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00192-020-04498-9>) contains supplementary material, which is available to authorized users.

✉ Mary H. Palmer
mhpalmer@email.unc.edu

¹ School of Nursing, University of North Carolina at Chapel Hill, CB 7460, Chapel Hill, NC, USA

² Division of Urology, Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Introduction

Approximately 2.3 billion of the adult world population experience lower urinary tract symptoms (LUTS), and of those, 1.7 billion are women with storage LUTS [1]. Storage LUTS, i.e., stress urinary incontinence (SUI), urgency urinary incontinence (UUI) and mixed urinary incontinence (MUI), urinary urgency, urinary frequency, and nocturia, have a detrimental impact on women’s health-related quality of life [1]. To improve women’s bladder health, an urgent need exists to disseminate evidence-based programs for storage LUTS to large groups or populations of women. This process, scaling-up of interventions, is complex. The World Health Organization (WHO) noted the core components for any scaling-up attempt, where an intervention’s effectiveness for representative samples and its scaling-up feasibility are considered [2].

Behavioral and pelvic floor muscle training (B-PFMT) programs are commonly used interventions to prevent and treat storage LUTS in women [3–5], specifically SUI. A core component of these programs is pelvic floor muscle exercise (PFME) with quick pelvic floor muscle contractions prior to an event triggering urine leakage, also known as the Knack, to prevent UI by inhibiting detrusor contraction [6, 7]. In women who have UUI and MUI, behavioral components including lifestyle modification and bladder training with urge suppression strategies are often combined with PFME as part of a multi-component B-PFMT program [6, 7]. These programs can be categorized as either supervised or unsupervised.

Supervised B-PFMT programs are conducted under conditions where women come to medical offices or clinics according to specified training intervals and participate in either individualized or group coached intervention that are typically provided by pelvic floor specialists (i.e., nurse specialists, physical therapists) [8]. A body of evidence exists for the effectiveness of supervised programs that aim to prevent and treat female storage LUTS in different age cohorts and stages of life [9–11]. Although supervised programs are highly recommended [12], they, by their nature, have limited feasibility to scale up to large groups or populations.

An inadequate number of qualified providers in clinical or community settings globally is a significant barrier to providing the intense level of supervision required by the program. As an example, the ratio of physical therapists to patients is < 1:1000 in Australia, the UK, the USA, and Canada, and the ratio is 1:100,000 in China [13]. The limited number of continence nurse specialists available in many countries is also concerning. Approximately 250 and 100 registered nurses specialize in continence nursing in Australia and Canada, respectively. Although 2302 nurses in Japan were certified as the wound, ostomy, and continence nurses in 2016, few had specialist knowledge and skills in the management of incontinence [14].

Besides workforce implications, women enrolled in supervised programs can face challenges. Because women need to travel to and from clinical locations, travel can act as a barrier to accessing care over time, especially for women living in rural areas. Frequent and long-distance transportations are reported as the barriers to sustaining exercise programs for individuals [15], and challenges women who live in rural areas face may lead to increased physical, psychological, and/or financial stress [16]. Because women are required to return for repeated visits to the setting where supervised programs are delivered, ancillary tasks of scheduling, preparing for, and following up from appointments can create additional work for the staff. When offering supervised programs to large numbers of women, dedicated and private space is needed, which can create logistic difficulties for clinic and community settings.

Unsupervised B-PFMT programs have been reported in the literature, and they are implicitly defined and typically reported to have two components: (1) provision of a single education session offered in face-to-face or non-face-to-face modalities to introduce participants to the programs and provide them with necessary information and materials and (2) participants' active self-administration of all aspects of the B-PFMT programs [17].

Because of the participants' independent role in these programs, they could avoid the aforementioned issues of feasibility posed by supervised programs. Moreover, unsupervised B-PFMT programs are acceptable to women. Qualitative evidence demonstrated that women who participated in unsupervised B-PFMT programs felt confident about self-training and thought it enabled them to assume responsibility for their symptom management [18]. Evidence of effectiveness is an important criterion for assessing the scalability of interventions. Therefore, we conducted this review to synthesize evidence of the effectiveness of unsupervised B-PFMT programs on improving storage LUTS outcomes including symptoms, severity, impact, self-reported symptom improvement, and pelvic floor muscle strength (PFMS) among adult community-dwelling women. Findings from this study may provide evidence for scaling up these programs in women living in the community.

Materials and methods

Search strategy

The systematic review was registered in PROSPERO (CRD42020149503). The report of this systematic review was guided by the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) statement [19]. Articles with women aged ≥ 18 years as participants were deemed eligible; other inclusion criteria for articles were: (1) ≥ 2 arm randomized controlled trials (RCTs) comparing effects of unsupervised B-PFMT intervention group(s) with control group(s) or with parallel intervention group(s); (2) quasi-experimental articles (i.e., articles using nonequivalent control group designs, pretest-posttest design, or interrupted time series design) reporting the effects of unsupervised B-PFMT programs; (3) short- and long-term outcomes relevant to storage LUTS, as defined by authors of retrieved studies, including symptoms, severity, impact, self-reported symptoms' improvement, and PFMS. Exclusion criteria for articles were: (1) case study/series, commentary, intervention protocol, and all type of review articles; (2) trials that combined B-PFMT programs with surgery or drug therapy; (3) women who were athletes, soldiers, described as frail, pregnant, and had cognitive impairment, multiple sclerosis, stroke, or lung disease; (4)

women who were performing biofeedback-assisted PFME or PFME using vaginal cones or electrical stimulation.

In consultation with a Health Sciences Library librarian, four databases—PubMed, CINAHL, Web of Science, and PsycINFO—were retrieved using search strings listed in Table S1. We searched the databases from their dates of inception through the last search date of August 6, 2019, and the language filter used for all databases was English.

Data extraction

The data extraction process was predominantly completed by two independent researchers. The titles and abstracts of the articles retrieved were assessed via Covidence (www.covidence.org) by rating the relevance of each article with “yes,” “maybe,” or “no” following the inclusion and exclusion criteria. The full texts for all articles rated as “yes” and “maybe” were further reviewed and assessed under the same criteria, and the final set of articles was determined by the reviewers. The data extraction form was developed by referring to the Data Collection Form for Intervention Review-Randomized Trials and Non-randomized Trials from the Cochrane Collaboration (<https://airways.cochrane.org/data-collection>). After pilot testing this form with eligible articles, it was then used to extract data. Any inconsistent rating arising between two researchers during the above steps was resolved by discussion and consensus. No direct contact with the authors of retrieved articles to gather additional or undisclosed information was made.

Risk of bias

The risk of bias for eligible articles was independently assessed by two researchers. The results were judged as “low,” “some concerns,” or “high” for RCTs by summarizing rating categories under five domains included in the tool of assessing risk of bias in randomized trials (RoB2) [20]. The risk of bias for quasi-experimental articles was evaluated as “the least risk of bias,” “some risk of bias,” or “significant risk of bias” by using the 12-item “Quality Assessment Tool for Before-After (Pre-Post) Articles With No Control Group” developed by the National Heart, Lung, and Blood Institute and Research Triangle Institute International (www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools). Two researchers discussed their ratings to reach consensus. If consensus could not be achieved, quality adjudication was forwarded to a methodologist external to the study whose expertise was in design and statistical methods.

Data synthesis

Variations existed in the: (1) methods used and providers recruited to deliver information or provide educational resource

materials to participants, (2) information delivered in the education session and included in the PFME elements, and (3) primary and secondary outcomes measures and their grading systems as well as analytic plans used. Because of these heterogeneities, pooling evidence to obtain an average number for effect size using meta-analysis was not applicable for this study [21]. Therefore, we used Popay et al.'s narrative synthesis approach, which demonstrated comparable synthesis power as meta-analysis to synthesize quantitative evidence in this review [22].

Results

Identification of articles

Our initial search strategy yielded 1388 articles including 368 duplicates across 4 databases. After title and abstract screening, 41 articles were moved to full-text screening. After the full-text screening, 13 articles remained eligible for this systematic review. Figure 1 depicts the selection process.

Summary of included articles

Table 1 outlines the overall characteristics of 13 eligible articles. Ten of the 13 articles were randomized controlled trials (RCTs) [23–32] and 3 used pretest-posttest designs [33–35]. Most of the studies (8/13) were from western countries, including the USA ($n = 4$) [23–25, 31] and Sweden ($n = 4$) [27–30]; three articles were from Turkey ($n = 2$) [26, 35] and Brazil ($n = 1$) [32], and two articles were from developed regions of China, i.e., Taiwan [33] and Hong Kong [34].

UI was the sole targeted storage LUTS identified in all the articles. In ten articles, the aim was to treat UI [24–30, 33–35], while the aim was to prevent UI in two articles [23, 31]. The aim of one article was to investigate the relationship between the frequency of PFME per day and pelvic floor muscle function [32]. In ten articles that reported UI treatment, women with unspecified UI, i.e., no UI type described, were enrolled in three articles [25, 33, 34], women with SUI were enrolled in four articles [27–30], women with either SUI or MUI were enrolled in two articles [26, 35], and in one article women with either SUI or UUI were enrolled [24]. In one article, women without pelvic floor muscle dysfunction were enrolled [32].

There were 2469 participants in the eligible articles. Women's ages ranged between 41 and 67 years in 12 articles and between 24 and 26 years old in 1 article [32]. The methodological evaluation demonstrated some concerns for bias in five of the ten RCTs [23–25, 31, 32] and high risk of bias for the remaining five RCTs [26–30]. Of three pretest-posttest articles, one article had the least risk of bias [33] and two had some risk of bias [34, 35].

Identification

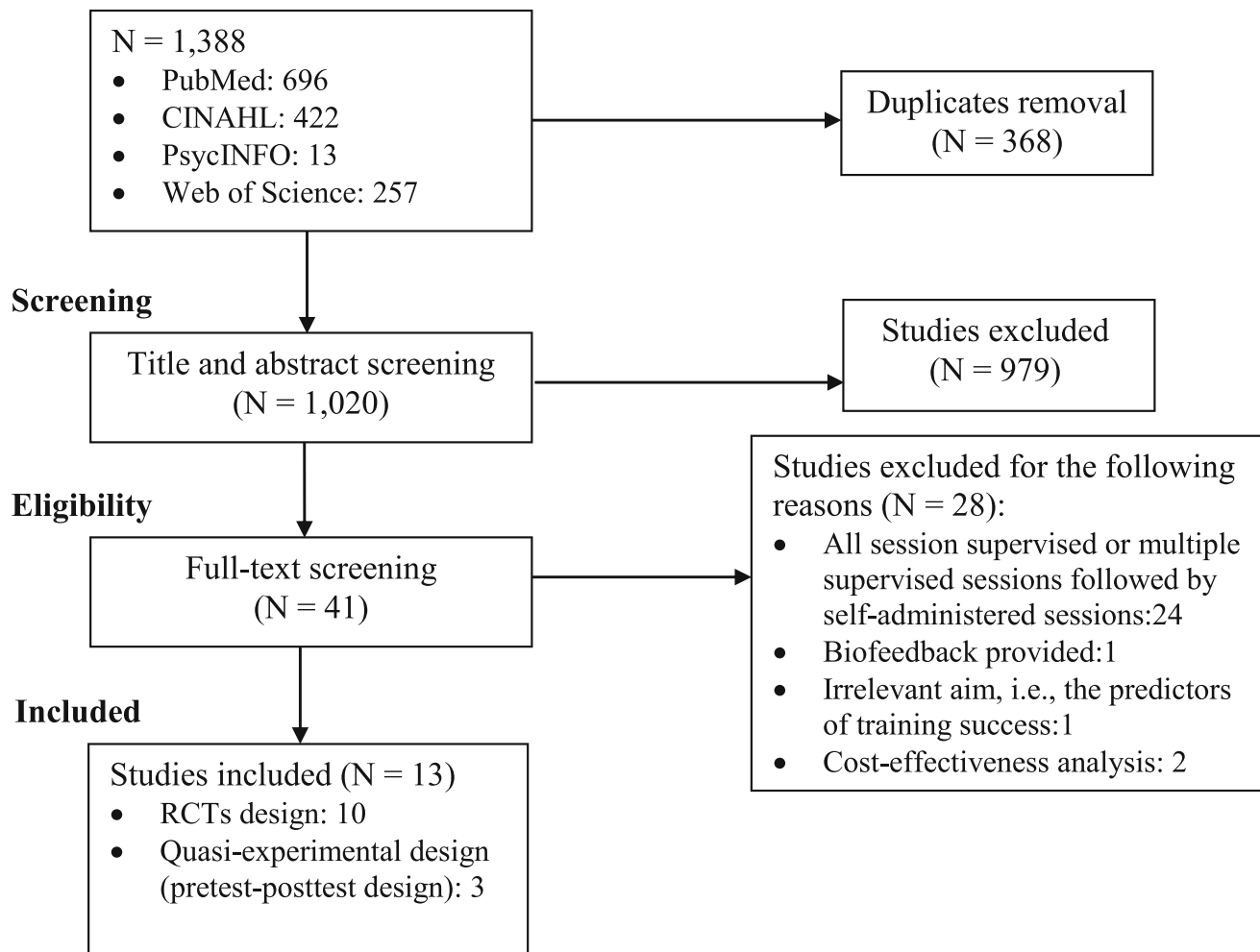


Fig. 1 PRISMA flow diagram for inclusion of articles in the systematic review

Components of unsupervised B-PFMT programs

Table S2 summarizes components of unsupervised B-PFMT programs reported in the eligible articles.

Education session

The method of delivery Eight of 13 articles described face-to-face education sessions including group delivery ($n = 4$) [23–25, 31] and one-on-one delivery ($n = 4$) [26, 32–34].

Five articles described non-face-to-face interactions including the use of emails, mailed materials, mobile Apps, and DVDs to deliver information [27–31], and one article did not report the method used to deliver information [35].

Provider(s) Those who provided face-to-face education sessions included a urologist ($n = 1$) [23], (trained) nurses specialists ($n = 3$) [23, 25, 31], physical therapists ($n = 2$) [32, 33], trained interventionist ($n = 1$) [24], and continence advisor

($n = 1$) [34]. One article did not describe background or discipline of the provider(s) [26].

Information provided PFME instructions were provided to all of the women; however, there was variation in the other information women received. Researchers in 12 articles reported that participants were taught how to locate and contract their pelvic floor muscle [23–30, 32–35]; in seven articles researchers used either vaginal palpation ($n = 6$) [23, 25, 26, 32, 33, 35] or having women draw-in their perineum or anus and contract their perineal muscles ($n = 1$) [35] as safeguards against incorrect practice. Ten articles reported that researchers provided information to increase participants UI knowledge [23–31, 33]; nine articles reported that researchers provided participants information about lifestyle modifications [23–25, 27–31, 34]; five articles reported that researchers provided participants with anatomical information for both the pelvis/pelvic floor and the lower urinary tract [23–26, 33]; three articles

Table 1 Overall characteristics of eligible articles ($n = 13$)

First author name, year, country	Design	Female participants	Risk of bias*
Diokno et al., 2004 USA	RCT	Having no urinary incontinence Intervention group: $n = 164$, age: 66.2 ± 6.4 years Control group: $n = 195$, age: 65.4 ± 6.7 years	Some concerns
Sarı et al., 2009 Turkey	RCT	Having stress or mixed urinary incontinence Intervention group: $n = 19$ (17 for analysis), age: 41.82 ± 8.65 years Control group: $n = 22$ (17 for analysis), age: 44.64 ± 6.90 years	High risk of bias
Diokno et al., 2010 USA	RCT	Having urinary incontinence Intervention group: $n = 23$, age: 60.6 ± 14.4 years Control group: $n = 21$ (18 for analysis), age: 52.2 ± 12.6 years	Some concerns
Hung et al., 2012 Taiwan, China	Pretest-posttest design	Having urinary incontinence ($n = 68$), age: 50.5 ± 6.0 years	Least risk of bias
Fan et al., 2013 Hongkong, China	Pretest-posttest design	Having urinary incontinence ($n = 372$), age: 52.3 ± 10.8 years	Some risk of bias
Sjöström, et al., 2013 Sweden	RCT	Having stress urinary incontinence Internet group: $n = 124$ (107 for analysis), age: 47.9 ± 10.6 years Postal group: $n = 126$ (113 for analysis), age: 49.4 ± 9.8 years	High risk of bias
Cavkaytar et al., 2015 Turkey	Pretest-posttest design	Having stress urinary incontinence ($n = 38$), age: 49.6 ± 8.1 years Having mixed urinary incontinence ($n = 34$), age: 48.9 ± 8.8 years	Some risk of bias
Sjöström et al., 2015 Sweden	RCT	Having stress urinary incontinence Internet group: $n = 124$ (87 for analysis at 1-year follow-up; 75 for analysis at 2-year follow-up), age: 47.9 ± 10.6 years Postal group: $n = 126$ (80 for analysis at 1-year follow-up; 79 for analysis at 2-year follow-up), age: 49.4 ± 9.8 years	High risk of bias
Asklund, et al., 2017 Sweden	RCT	Having stress urinary incontinence Mobile App group: $n = 62$ (61 for analysis), age: 44.8 ± 9.7 years Control group: $n = 61$ (60 for analysis), age: 44.7 ± 9.1 years	High risk of bias
Hoffman et al., 2017 Sweden	RCT	Having stress urinary incontinence Mobile App group: $n = 62$ (46 for analysis at follow-up), age: 44.8 ± 9.7 years	High risk of bias**
Sampsel et al., 2017 USA	RCT	Having no urinary incontinence 2-h class group: $n = 332$ (298 for analysis at 3-month follow-up; 291 for analysis at 12-month follow-up; 276 for analysis at 24-month follow-up), age: 63.03 ± 5.43 years 20-min DVD group: $n = 315$ (290 for analysis at 3-month follow-up; 280 for analysis at 12-month follow-up; 268 for analysis at 24-month follow-up), age: 62.79 ± 5.91 years	Some concerns
Diokno, et al., 2018 USA	RCT	Having stress or urgency urinary incontinence Intervention group: $n = 232$ (209 for analysis at 3-month follow-up; 192 for analysis at 6-month follow-up; 184 for analysis at 9-month follow-up; 195 for analysis at 12-month follow-up), age: 64 ± 7 years Control group: $n = 231$ (212 for analysis at 3-month follow-up; 205 for analysis at 6-month follow-up; 202 for analysis at 9-month follow-up; 203 for analysis at 12-month follow-up), age: 65 ± 8 years	Some concerns
Pereira-Baldon, et al., 2019 Brazil	RCT	Having no reported pelvic floor muscle dysfunction Intervention group 1 (frequency: once daily): $n = 15$ (13 for analysis), age: 25.08 ± 2.79 years Intervention group 2 (frequency: 3 times daily): $n = 15$ (12 for analysis), age: 24.26 ± 3.82 years	Some concerns

Notes

(1) RCT: randomized controlled trial

(2) *Risk of bias for an RCT was judged as “low risk of bias,” “some concerns,” and “high risk of bias” by using the Cochrane tool of assessing risk of bias in randomized trials (RoB2), while risk of bias for pretest-posttest design was graded on “least risk of bias,” “some risk of bias,” and “significant risk of bias” by using 12-item tool of “Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group” developed by the National Heart, Lung, and Blood Institute and Research Triangle Institute International

(3) **Although the article only used the intervention group to answer research aim(s), data were derived from the RCT reported by Asklund et al., 2017; therefore, the design and risk of bias were assessed based on this information

reported that researchers provided participants anatomical information about the pelvis/pelvic floor only [29–31]. Two articles reported researchers taught participants the “Knack” and bladder training [24, 31], five articles reported researchers taught participants the “Knack” only [26–30], and two articles reported researchers taught participants the bladder training only [23, 25]. One article reported researchers provided participants with information about neural control of the lower urinary tract [23].

Self-administered training

Elements of PFME Elements of PFME were reported in 9 of 13 articles [26–30, 32–35] and included repetition, frequency of exercises, and duration of exercises. Information covered by these elements varied across articles except for the frequency of exercises: three sets per day were reported in seven articles [26–30, 33, 34].

Reinforcement strategies Ten articles reported the reinforcement strategies researchers used in their programs [23–30, 34, 35]. Specifically, four articles reported that face-to-face contacts with participants were used for re-assessing the correct contraction of pelvic floor muscles at 10 days ($n = 1$) [35], at 2 to 4 weeks ($n = 2$) [23, 25], at 3 months, and 6 to 9 months ($n = 1$) [34] after the intervention initiation. Researchers in four articles applied strategies to promote adherence [24, 26, 29, 30], by either providing weekly telephone contacts or a magnetized reminder that displayed the project logo to serve as a discrete reminder to follow the program ($n = 2$) [24, 26], or contacting participants through an email at 4 weeks, or allowing participants to create three reminders per day in a mobile App after the intervention initiation ($n = 2$) [29, 30]. Two articles reported that researchers provided participants timely support and answered questions initiated by participants through email [27, 28], and one study reported that researchers initiated telephone contact with the participants on a weekly basis to answer questions participants raised [26].

Outcome assessment tools

Table 2 describes all the measures used to assess outcomes.

- (1) Symptoms diagnostic/screening tools ($n = 8$) included 2- or 3-day bladder diary ($n = 7$) [23–27, 29, 33], 1-h or 24-h pad test ($n = 3$) [24–26], and paper towel test ($n = 3$) [24, 25, 31]; two articles included all three tools [24, 25], and one article included the first two tools [26].
- (2) Symptom severity assessment tools ($n = 9$) included the standardized Medical Epidemiologic and Social aspects of Aging (MESA) questionnaire ($n = 2$) [23, 24], the Sandvik Severity Index ($n = 1$) [25], Severity Index Score ($n = 1$) [33], the Indevus Urgency Severity Scale

(IUSS) ($n = 1$) [31], and the International Consultation on Incontinence Questionnaire-Urinary Incontinence-Short Form (ICIQ-UI SF) ($n = 6$) [24, 27–31]. Two articles included two tools, i.e., the IUSS and ICIQ-UI SF [31] and MESA and ICIQ-UI SF [24], respectively.

- (3) PFMS assessment tools ($n = 7$) included digital palpation ($n = 6$) [23–25, 32, 33, 35] and pressure perineometer ($n = 1$) [26].
- (4) Perceived symptom improvement assessment tools ($n = 7$) included one self-reported improvement question ($n = 1$) [33] and the Patient Global Impression of Improvement (PGI-I) ($n = 6$) [24, 27–30, 35].
- (5) Symptom impacts assessment tools ($n = 9$) included the Symptom Impact Index ($n = 1$) [33], the Incontinence of Quality of Life (I-QOL) ($n = 2$) [24, 26], the Urogenital Distress Inventory-6 (UDI-6) ($n = 2$) [34, 35], the Incontinence Impact Questionnaire short form (IIQ-7) ($n = 2$) [34, 35], the International Consultation on Incontinence Questionnaire-Lower Urinary Tract Symptoms Quality of Life (ICIQ-LUTSqol) ($n = 4$) [27–30], and the EuroQol 5D-Visual Analogue Scale (EQ5D-VAS) ($n = 2$) [27, 28]. Two articles included both the UDI-6 and IIQ-7 [34, 35], and two articles included both the ICIQ-LUTSqol and EQ5D-VAS [27, 28].

Outcome synthesis of studies

Table S3 describes outcomes assessed by symptom diagnosis/screening tools.

Bladder diary

UI treatment Two articles reported significant reduction of number of voids for the intervention group (at least 6–8 weeks post intervention) compared to the control group [24, 25]. Three articles reported the significant reduction of the number of UI episodes after at least 2 months’ intervention for the intervention group compared to the control group [24, 26, 29]; one article reported a comparably significant reduction of the number of UI episodes for two parallel groups (i.e., an internet intervention administered group and a postal intervention administered group) after 4 months’ intervention [27]. One pretest-posttest article reported significant reduction in the number of voids and the number of UI episodes after 4 months’ intervention [33].

UI prevention One article reported that continent participants had a significant reduction of the number of voids at 12 months after intervention for the intervention group compared to the control group [23].

Table 2 Tools for assessing outcomes of interest (*n* = 13)

First author name, year, country	Symptom diagnostic/screening tools		Symptom severity assessment tools		Pelvic floor muscle strength assessment tools		Perceived symptom improvement assessment tools		Symptom impact assessment tools				
	1-h/24-h pad test	Paper towel test	MESA Severity Index	Severity Index Score	IUSS UI SF	Digital palpation	Pressure perineometer	Self-reported improvement	PGI-I	Symptom Impact Index	I-QOL 6	IIQ-7	ICIQ-LUTSqol VAS
Diokno et al., 2004, USA	✓		✓			✓							
Sari et al., 2009, Turkey	✓*						✓				✓		
Diokno et al., 2010, USA	✓	✓		✓		✓							
Hung et al., 2012, Taiwan, China	✓			✓		✓		✓					
Fan et al., 2013, Hong Kong, China											✓	✓	
Sjöström et al., 2013, Sweden	✓**				✓				✓				✓
Cavkaytar et al., 2015, Turkey						✓			✓				
Sjöström et al., 2015, Sweden					✓				✓				✓
Asklund et al., 2017, Sweden	✓**				✓				✓				✓
Hoffman et al., 2017, Sweden					✓				✓				✓
Sampselle et al., 2017, USA		✓											
Diokno et al., 2018, USA	✓	✓				✓							
Pereira-Baldon et al., 2019, Brazil											✓		

Notes

(1) * 1-h pad test

(2) ** 2-day bladder diary

(3) Abbreviations

MESA: the standardized Medical Epidemiologic and Social aspects of Aging questionnaire

IUSS: Indevus Urgency Severity Scale

ICIQ-UI SF: International Consultation on Incontinence Questionnaire-Urinary Incontinence-Short Form

PGI-I: Patient Global Impression of Improvement

I-QOL: Incontinence of Quality of Life

UDI-6: Urogenital Distress Inventory

IIQ-7: Incontinence Impact Questionnaire short form

ICIQ-LUTSqol: International Consultation on Incontinence Questionnaire-Lower Urinary Tract Symptoms Quality of Life

EQ5D-VAS: EuroQol 5D-Visual Analogue Scale

Pad test

UI treatment Two articles reported no significant advantages of the intervention group over the control group on urine leakage reduction in grams assessed by 24-h pad test after 6 to 8 weeks' intervention, but significantly fewer grams of urine leakage in the intervention group versus that in the control group at 3 months and 12 months [24, 25]. One article reported a significantly greater reduction in grams of urinary leakage assessed by the 1-h pad test for the intervention group than that in the control group from baseline to 2 months [26].

Paper towel test

UI treatment Two articles reported a significant reduction in mean leak diameter (i.e., the sum of orthogonal diameters of the wet area divided by two) after 6 to 8 weeks' intervention or lower percentage of participants having a positive paper towel test at 3 months and 12 months in the intervention group versus that in the control group [24, 25].

UI prevention One article enrolled continent participants in two parallel groups (i.e., a class intervention administered group and a DVD intervention administered group) and reported neglectable changes in the paper towel test results between baseline and each of three follow-up time points for each group [31].

Table S4 provides outcomes measured by symptom severity assessment tools.

ICIQ-UI SF

UI treatment Four articles reported significant reductions in post-intervention scores, with the mean differences (MD) ranging from 2.9 to 3.9 from baseline (with the score rated 10 and above) to 3 months, 12 months and 24 months [27–30], but the differences in the intervention effect across time points were not presented by these data. One article reported that the ICIQ-UI SF scores decreased 1.96 points on average every 3 months for the intervention group, which was significantly larger than that for the control group; the average reduction was 0.98 [24]. Two articles reported comparably significant reductions in scores between two parallel groups (i.e., an internet intervention administered group and a postal intervention administered group) of participants at the following measurement intervals: from baseline to 3 months, 12 months, and 24 months [27, 28].

UI prevention One article reported comparable reductions in scores between two parallel groups (i.e., a class intervention administered group and a DVD intervention administered group) of participants at the following measurement intervals: baseline to 3 months, 12 months, and 24 months [31].

Mesa

UI treatment For participants who had SUI and those who had UII, one article reported the median sum scores of all items of MESA were significantly lower for the intervention group than those for the control group at 3 months and 12 months [24].

UI prevention Instead of addressing all items in MESA, researchers in one article enrolled continent participants and operationalized continence as both having no leakage and having leakages no more than 5 days in the past 12 months. They reported the odds of having no leakage at 12 months for participants in the intervention group was 2.03 times (95% CI 1.04–3.98, $p = 0.04$) that for participants in the control group. They also reported that the odds of continence status remaining unchanged and transitioning from no more than 5 days to no leakage from baseline to 12 months for participants in the intervention group was 1.97 times (95% CI 1.15–3.98, $p = 0.01$) that for participants in the control group [23].

Other severity tools

UI treatment One article used the Sandvik Severity Index to classify UI into three severity categories, i.e., slight, moderate, and severe based on the frequency and amount of urine leakage at baseline and 6 to 8 weeks post-intervention. There was a significant decrease in the percentage of participants in the moderate cluster (47.8% to 21.7%, $p = 0.03$), and there was a significant increase in the percentage of participants in the slight cluster (17.4% to 56.5%, $p = 0.036$). There were no significant changes for participants in the control group in each severity category [25]. One article reported a significant decrease in UI severity assessed by Severity Index Score after a 4-month intervention, i.e., the median score changed from six at baseline to three at 4 months, with $p < 0.001$ [33].

UI prevention One article reported there was no significantly different amelioration in urinary urgency severity assessed by IUSS between two parallel groups (i.e., class intervention administered group and the DVD intervention administered group) of participants from baseline to 3 months, 12 months, and 24 months [31].

Table S5 describes PFMS assessed by digital palpation and pressure perineometer.

Digital palpation

UI treatment Using grading on the Brink scoring system, one article reported a significant increase in scores for pressure, displacement, and duration after 6 to 8 weeks of intervention, while in the control group, a significant increase was found for displacement [25]. One article reported no significant

differences between the intervention group and the control group at baseline, 3 months, and 12 months in percentages of participants who were graded 4, 5, or 6 for pressure and who were graded 4 or 5 for displacement and in median scores for duration [24]. Using grading on the Modified Oxford Scale, two pretest-posttest articles reported significant increases in PFMS at 2 and 4 months post-intervention [33, 35].

UI prevention Using grading on the Brink scoring system, one article reported there were significantly higher scores of pressure and displacement at 12 months and significantly higher increases in these scores from baseline to 12 months for continent participants in the intervention group versus those in the control group [23]. Using grading on the Modified Oxford Scale, one article reported a significant increase in PFMS for participants doing PFME once daily and those doing PFME three times daily when assessed at 2 months after the intervention initiation. There were no significant differences in PFMS between the two groups at baseline and at 2 months [32].

Pressure perineometer

UI treatment One article reported the increases in the mean contraction pressure and maximum contraction pressure of pelvic floor muscle were significantly greater for participants in the intervention group than for those in the control group [26].

Table S6 describes outcomes assessed by perceived symptom improvement assessment tools.

PGI-I

UI treatment Grading on a 7-point Likert scale from “very much better” to “very much worse,” two articles reported that the percentages of participants with their UI getting much better or very much better were significantly higher in the internet group than those in the postal group at 4 months (40.9% versus 26.5%, $p = 0.01$) and at 24 months (39.2% versus 23.8%, $p = 0.03$), but the significant difference was not observed at 12 months [27, 28]. One article reported that significantly more participants in the intervention group said their UI was much better or very much better than those in the control group [29]. One article reported the percentage of participants who said that their UI was much better or very much better was significantly higher in the intervention group than that in the control group at 3 months (46.9% versus 8.1%, $p < 0.001$) and at 12 months (64.3% versus 11.3%, $p < 0.001$) [24]. Another article reported that 66.7% of participants in the intervention group reported their leakages were much better or very much better assessed at 24 months [30]. Grading on yes/no improvement responses, an article reported findings from a pretest-posttest study in which the percentage of participants with SUI who graded on “yes” was

significantly higher than those with MUI (68.4% versus 41.2%, $p = 0.01$) at 2 months [35].

One self-reported improvement question

UI treatment One pretest-posttest article reported that 75% of participants reported their UI was “improved” and “cured” at 4 months [33].

Table S7 describes outcomes evaluated by symptom impacts assessment tools.

ICIQ-LUTSqol

UI treatment Two articles reported significant reductions in scores for both the internet group and postal group from baseline to 4 months ($MD_{\text{internet}} = 5.8$; $MD_{\text{postal}} = 4.8$), to 12 months ($MD_{\text{internet}} = 6.1$; $MD_{\text{postal}} = 5.8$), and to 24 months ($MD_{\text{internet}} = 7.1$; $MD_{\text{postal}} = 6.4$) [27, 28], but there were no significant differences in reductions between groups; two articles reported significant reductions in scores for the intervention group from baseline to 3 months ($MD = 4.8$) and to 24 months ($MD = 4.0$), and participants in the intervention group had a significantly lower score than those in the control group at 3 months [29, 30].

I-QOL

UI treatment One article reported the increases of total scores and scores for each of three domains (i.e., avoidance and limiting behavior, psychosocial impacts, and social embarrassment) were significantly higher for the intervention group than those for the control group from baseline to 2 months (23.19 ± 11.43 versus -5.74 ± 6.26 , $p < 0.01$) [26]. One article reported the total scores were significantly higher for the intervention group than those for the control group at 3 months (median: 86 versus 83, $p < 0.001$) and at 12 months (median: 92 versus 85, $p < 0.001$) [24].

UDI-6 and IIQ-7

UI treatment Two pretest-posttest articles reported a significant reduction in UDI-6 and IIQ-7 scores for participants with UI ($MD_{\text{UDI-6}} = 8.6$, $MD_{\text{IIQ-7}} = 7.3$), with SUI ($MD_{\text{UDI-6}} = 26.1$, $MD_{\text{IIQ-7}} = 21.9$), and with MUI ($MD_{\text{UDI-6}} = 13.1$, $MD_{\text{IIQ-7}} = 15.2$) [33, 34]. One of them also reported the reduction in UDI-6 and IIQ-7 scores were significantly larger for participants with SUI than for those with MUI [35].

EQ5D-VAS

UI treatment Two articles reported a significant increase in scores for participants in the internet group from baseline to 4 months ($MD = 4.2$) and to 24 months ($MD = 4.2$), but there

were no significant differences in score increases between participants in the internet group and those in the postal group [27, 28].

Symptom impact index

UI treatment One article reported significant reductions in scores for four items (i.e., the number of worries, the number of activities affected, avoiding activities because of worrying about leakages, and avoiding activities because of needing a toilet) from baseline to 4 months [33].

Discussion

This review provides evidence that unsupervised B-PFMT programs for middle-aged women who have UI are appropriate for scaling up to the population level. With the high prevalence and impact burden of storage LUTS, especially UI, efforts to provide population-based interventions are needed. Synthesized evidence resulting from this study identifies characteristics of women most often studied, unpacks unsupervised B-PFMT programs into their components, describes outcome assessment modules, and provides accumulated evidence supporting the effectiveness of these programs on treating women's UI. This evidence also indicates that unsupervised B-PFMT programs appear to be a promising scaling-up approach while providing important guidance for scaling-up attempts with unsupervised B-PFMT programs.

Women with UI represented the majority of participants in the eligible articles ($n = 10$) and were mostly middle-aged (i.e., 40 to 60 years old); three articles describe prevention-focused unsupervised B-PFMT programs, i.e., women who did not have UI ($n = 2$) or pelvic floor muscle dysfunction ($n = 1$). Few articles were located that primarily enrolled women < 40 or > 60 years old. Unsupervised B-PFMT programs have not been tested in the prevention or treatment of storage LUTS other than UI. These include nocturia, urinary urgency, and urinary frequency in women across the life course. Future studies are recommended to address these gaps to advance the science of preventing and treating storage LUTS among women.

Unsupervised B-PFMT programs are conceptually defined as having a one-time education session followed by a long-term self-administered training program. There are, however, variations in how some researchers operationalize such programs. First, multiple modalities for information delivery were used in the education session, including group and individual delivered, face-to-face and non-face-to-face delivered (i.e., mailing materials and adopting DVD, internet, and mobile Apps). Second, although the information delivered in education sessions generally adhered to UI conservative behavioral management guidelines, including PFME, bladder

training, and lifestyle modification, variations exist with the inclusion of other information, e.g., teaching information about female anatomy of the lower urinary tract, UI, and/or nervous system controlling the lower urinary tract. It remains unknown if the type of information delivered in an education session influences the quality and quantity of subsequent self-administered training. Third, information about the elements of PFME in the published articles included repetition, frequency of exercises, and duration of exercise. Except for exercise frequency in the form of three sets per day, information for the other elements differed dramatically across articles. This variation makes replicating and building on research findings challenging. The use of checklists in publications, such as consensus on the exercise-reporting template (CERT) [36] and template for intervention description and replication (TIDieR) [37], is recommended for future studies.

Unfortunately, little information was reported about participants' practice of lifestyle modification, bladder training, urge suppression strategies, and the "Knack" during self-administered training; thus, the magnitude of the effects of these behavioral components on storage LUTS are underexplored. Monitoring participants' performance and adherence to these behavioral components and testing their effects on outcomes are recommended before scaling-up attempts.

Another observation from this review is the lack of core outcomes and core measurement tools. Researchers used multiple assessment tools in an attempt to capture parameters indicating UI changes. They can be categorized into the following modules: symptom diagnostic/screening tools, symptom severity assessment tools, perceived symptom improvement tools, and PFMS assessment tools.

Researchers also used either symptom-specific or generic symptom impact assessment tools to quantify the changes in quality of life and disturbances of UI on individuals' activities, relationships, and feelings after the intervention. Two methodological strategies are recommended for future scaling-up programs. First, careful selection of tools from each assessment module is needed by giving comprehensive consideration of their relevance to participants in the study (i.e., tools used for individuals with specific symptoms or without specific symptoms) [38], their psychometric characteristics (i.e., reliability and validity tested under the classic test theory or difficulty and discrimination tested under the item response theory) [39], participants' characteristics, which might influence their understanding (e.g., literacy), and the feasibility of application to large groups or populations. Second, preplanning approaches to adjust the p value for multiple comparisons to avoid the inflation of the type I error and monitoring data presentation to avoid selective reporting and p-hacking are also important [40].

This review found the amelioration of UI symptoms, severity, impact, subjective improvement of UI, and

improvement of PFMS were evident at 6 to 8 weeks after program initiation. Cumulative effectiveness however was limited to specific outcomes, i.e., reduction in the number of UI episodes as evidence from bladder diary entries, reduction of ICIQ-UI SF and ICIQ-LUTqol scores, improvement of symptoms assessed by PGI-I, and improvement of PFMS assessed by vaginal palpation.

The reduction in the number of UI episodes after \geq 2 months was evident from this review despite the various descriptive statistics reported in articles, i.e., percentage, median, and mean. Conclusions about the effect size and its change over time remain limited.

The significant reduction in ICIQ-SF scores after 3 months or longer, indicating improvement of UI, was evident from this review, but the effect size described in mean difference might not be influenced by the time variable. This change reflects clinically meaningful differences given the significant improvement of ICIQ-LUTS_{qol} and subjective perception of symptom improvement assessed by PGI-I, which were collected from women with their ICIQ-SF responses. It remains unclear however if the effect size for objective and subjective improvement could be influenced by using different combinations of the unsupervised B-PFMT program components.

As the only ‘sign,’ significant improvement of PFMS was evident 6 to 8 weeks after starting the programs, and this finding was not altered by the type of statistic (i.e., median, percentage and mean) or grading systems used. Neither the pooled effect size of this outcome nor its changes over time can be concluded from this review. In addition, this outcome cannot be assessed without face-to-face contact with women, which limits its application in scaling-up programs.

Despite the promising findings of applying unsupervised B-PFMT programs to prevent UI among postmenopausal women, no cumulative effectiveness can be obtained in this review. More UI prevention studies of women across the life course are warranted to determine their effectiveness in promoting bladder health and inclusion in scaling-up efforts.

Limitations

This systematic review has several limitations. There were concerns about quality (i.e., 5 RCTs had high risk of bias, 2 pretest-posttest articles had some risk of bias) of eligible articles in this review may compromise some conclusions we made. Rigorous studies applying unsupervised B-PFMT programs are required to provide a high level of evidence. No contact with the authors in retrieved articles was made during this study. Therefore, some detailed information about the intervention protocols and findings may not have been included in this review. Effect sizes for most significant findings cannot be synthesized from

this review, and it remains unclear if they are influenced by time or various combinations in the components of the unsupervised B-PFMT programs. Specific populations, e.g., pregnant and postpartum women, were not represented in this review because variance of data from these groups could compromise the precision of effectiveness synthesis for the majority of women. Therefore, the findings from this review can be extrapolated only to UI treatment of women in their 40s to 60s who live in the community. Initial scaling-up attempts may have to be situated within this limitation.

Conclusions

Evidence from this review indicates that unsupervised B-PFMT programs can be scaled up to women in their 40s to 60s who have UI. More studies are needed across the life course to investigate potential effects of unsupervised B-PFMT programs with women who do not have UI and with women who have storage LUTS other than UI.

No optimal composition of unsupervised B-PFMT programs can be concluded from this review, but researchers can use this information to address the identified gaps in knowledge. Unsupervised B-PFMT programs have the potential to be scaled up to improve women’s access to B-PFMT programs and improve their bladder health.

Acknowledgements We are grateful to Jamie Conklin, MLIS, for helping us retrieve articles from databases and to George Knafl, PhD, for providing methodological consultation for this study.

Authors’ contributions C Wu, MH Palmer, and DK Newman: Protocol development.

C Wu, MH Palmer: Data management, data analysis, manuscript writing, manuscript editing.

DK Newman: Manuscript editing.

Compliance with ethical standards

Conflict of interests None.

Abbreviations LUTS, Lower urinary tract symptoms; UI, Urinary incontinence; SUI, Stress urinary incontinence; UUI, Urgency urinary incontinence; MUI, Mixed urinary incontinence; PFMT, Pelvic floor muscle training; PFME, Pelvic floor muscle exercise; RCT, Randomized controlled trial; PFMS, Pelvic floor muscle strength; MESA, The standardized Medical Epidemiologic and Social Aspects of Aging questionnaire; IUSS, Indevus Urgency Severity Scale; ICIQ-UI SF, The International Consultation on Incontinence Questionnaire-Urinary Incontinence-Short Form; PGI-I, The Patient Global Impression of Improvement; I-QOL, Incontinence of Quality of Life; UDI-6, Urogenital Distress Inventory-6; IIQ-7, Incontinence Impact Questionnaire short form; ICIQ-LUTS_{qol}, International Consultation on Incontinence Questionnaire-Lower Urinary Tract Symptoms Quality of Life; EQ5D-VAS, EuroQol 5D-Visual Analogue Scale

References

- Abrams P, Cardozo LD, Wagg A, Wein AJ (2017) International consultation on incontinence. 6th ed. Plymouth, UK: Health Publications Ltd.
- World Health Organization and Department of Reproductive Health and Research-ExpandNet (2010) Nine steps for developing a scaling-up strategy. https://www.who.int/reproductivehealth/publications/strategic_approach/9789241500319/en/.
- Newman DK, Wein AJ. Office-based behavioral therapy for management of incontinence and other pelvic disorders. *Urol Clin*. 2013;40(4):613–35. <https://doi.org/10.1016/j.ucl.2013.07.010>.
- Newman DK, Wein AJ (2008) Managing and treating urinary incontinence. 2nd ed. Health Professions Press.
- Lamin E, Parrillo LM, Newman DK, Smith AL. Pelvic floor muscle training: underutilization in the USA. *Curr Urol Rep*. 2016;17(2):10. <https://doi.org/10.1007/s11934-015-0572-0>.
- Newman DK, Borello-France D, Sung VW. Structured behavioral treatment research protocol for women with mixed urinary incontinence and overactive bladder symptoms. *Neurourol Urodyn*. 2018;37(1):14–26. <https://doi.org/10.1002/nau.23244>.
- Newman DK, Burgio K (2016) Conservative management of urinary incontinence: behavioral and pelvic floor therapy and urethral and pelvic devices. Campbell-Walsh urology. 11th ed. Philadelphia, PA: Elsevier.
- 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. *Eur J Obstet Gynecol Reprod Biol*. 2011;159(2):465–71. <https://doi.org/10.1016/j.ejogrb.2011.09.003>.
- Palmer MH, Cockerell R, Griebing TL, Rantell A, van Houten P, Newman DK. Review of the 6th international consultation on incontinence: primary prevention of urinary incontinence. *Neurourol Urodyn*. 2020;39(1):66–72. <https://doi.org/10.1002/nau.24222>.
- Cacciari LP, Dumoulin C, Hay-Smith EJ. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women: a Cochrane systematic review abridged republication. *Braz J Phys Ther*. 2019;23(2):93–107. <https://doi.org/10.1016/j.bjpt.2019.01.002>.
- Monteiro S, Riccetto C, Araújo A, Galo L, Brito N, Botelho S. Efficacy of pelvic floor muscle training in women with overactive bladder syndrome: a systematic review. *Int Urogynecol J*. 2018;29(11):1565–73. <https://doi.org/10.1007/s00192-018-3602-x>.
- Bo K, Frawley HC, 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. *Int Urogynecol J*. 2017;28(2):191–213. <https://doi.org/10.1007/s00192-016-3123-4>.
- Jones A, Skinner MA. The current status of physical therapy in China. *Chinese J Rehabil Med*. 2013;28(6):493–501.
- Ostaszkiwicz J, Peden-McAlpine C, Northwood M, Eustice S, Bliss DZ, Nishimura K (2018) Advanced Practice Continence Nursing. In *Management of Fecal Incontinence for the Advanced Practice Nurse* (pp. 15–47). Springer, Cham.
- Valenzuela T, Okubo Y, Woodbury A, Lord SR, Delbaere K. Adherence to technology-based exercise programs in older adults: a systematic review. *J Geriatr Phys Ther*. 2018;41(1):49–61. <https://doi.org/10.1519/JPT.0000000000000095>.
- Field M (2019) Power and injustice: exploring the factors that limit Women's access to maternal health Care in rural Pennsylvania. Dissertation, Pennsylvania State University.
- Newman DK, Wein AJ. Office-based behavioral therapy for management of incontinence and other pelvic disorders. *Urol Clin North Am*. 2013;40(4):613–35. <https://doi.org/10.1016/j.ucl.2013.07.010>.
- Asklund I, Samuelsson E, Hamberg K, Umefjord G, Sjöström M. User experience of an app-based treatment for stress urinary incontinence: qualitative interview study. *J Med Internet Res*. 2019;21(3):e11296. <https://doi.org/10.2196/11296>.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>.
- Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. <https://doi.org/10.1136/bmj.l4898>.
- Fletcher J. What is heterogeneity and is it important? *BMJ*. 2007;334(7584):94–6. <https://doi.org/10.1136/bmj.39057.406644.68>.
- Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, Britten N, Roen K, Duffy S (2006) Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version 1:b92.
- Diokno AC, Sampsel CM, Herzog AR, Raghunathan TE, Hines S, Messer KL et al. Prevention of urinary incontinence by behavioral modification program: a randomized, controlled trial among older women in the community. *J Urol*. 2004;171(3):1165–71. <https://doi.org/10.1097/01.ju.0000111503.73803.c2>.
- Diokno AC, Newman DK, Low LK, Griebing TL, Maddens ME, Goode PS et al. Effect of group-administered behavioral treatment on urinary incontinence in older women: a randomized clinical trial. *JAMA Intern Med*. 2018;178(10):1333–41. <https://doi.org/10.1001/jamainternmed.2018.3766>.
- Diokno AC, Ocampo MS, 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. *Int Urol Nephrol*. 2010;42(2):375–81. <https://doi.org/10.1007/s11255-009-9626-x>.
- Sar D, Khorshid L. The effects of pelvic floor muscle training on stress and mixed urinary incontinence and quality of life. *J Wound Ostomy Continence Nurs*. 2009;36(4):429–35. <https://doi.org/10.1097/WON.0b013e3181aaf539>.
- Sjöström 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 Int*. 2013;112(3):362–72. <https://doi.org/10.1111/j.1464-410X.2012.11713.x>.
- 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 Int*. 2015;116(6):955–64. <https://doi.org/10.1111/bju.13091>.
- Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a randomized controlled trial. *Neurourol Urodyn*. 2017;36(5):1369–76. <https://doi.org/10.1002/nau.23116>.
- Hoffman V, Söderström L, Samuelsson E. Self-management of stress urinary incontinence via a mobile app: two-year follow-up of a randomized controlled trial. *Acta Obstet Gynecol Scand*. 2017;96(10):1180–7. <https://doi.org/10.1111/aogs.13192>.
- Sampsel CM, Newman DK, Miller JM, Kirk K, DiCamillo MA, Wagner TH, et al. A randomized controlled trial to compare 2 scalable interventions for lower urinary tract symptom prevention: Main outcomes of the TULIP study. *J Urol*. 2017;197(6):1480–6. <https://doi.org/10.1016/j.juro.2016.12.099>.
- Pereira-Baldon VS, Avila MA, Dalarmi CB, de Oliveira AB, Driusso P. Effects of different regimens for pelvic floor muscle

- training in young continent women: randomized controlled clinical trial. *J Electromyogr Kinesiol.* 2019;44:31–5. <https://doi.org/10.1016/j.jelekin.2018.11.008>.
33. Hung HC, Chih SY, Lin HH, Tsauo JY. Exercise adherence to pelvic floor muscle strengthening is not a significant predictor of symptom reduction for women with urinary incontinence. *Arch Phys Med Rehabil.* 2012;93(10):1795–800. <https://doi.org/10.1016/j.apmr.2012.03.010>.
 34. Fan HL, Chan SS, Law TS, Cheung RY, Chung TK. Pelvic floor muscle training improves quality of life of women with urinary incontinence: a prospective study. *Aust N Z J Obstet Gynaecol.* 2013;53(3):298–304. <https://doi.org/10.1111/ajo.12075>.
 35. Cavkaytar S, Kokanali MK, Topcu HO, Aksakal OS, Doğanay M. Effect of home-based Kegel exercises on quality of life in women with stress and mixed urinary incontinence. *J Obstet Gynaecol.* 2015;35(4):407–10. <https://doi.org/10.3109/01443615.2014.960831>.
 36. Slade SC, Dionne CE, Underwood M, Buchbinder R. Consensus on exercise reporting template (CERT): explanation and elaboration statement. *Br J Sports Med.* 2016;50(23):1428–37. <https://doi.org/10.1136/bjsports-2016-096651>.
 37. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ.* 2014;348:g1687. <https://doi.org/10.1136/bmj.g1687>.
 38. Sussman RD, Syan R, Brucker BM. Guidelines of the guidelines: urinary incontinence in women. *BJU Int.* 2020;125(5):638–55. <https://doi.org/10.1111/bju.14927>.
 39. DeVellis RF (2016) Scale development: theory and applications. 4th ed. Sage publications.
 40. Vidgen B, Yasseri T. P-values: misunderstood and misused. *Front Phys.* 2016;4(6). <https://doi.org/10.3389/fphy.2016.00006>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.