National Institute for Health and Care Excellence

Draft for consultation

Urinary incontinence (update) and pelvic organ prolapse in women: management

NICE guideline tbc

Supplementary material C

October 2018

Methods

Draft for Consultation

Supplementary material was developed by the National Guideline Alliance, hosted by the Royal College of Obstetricians and Gynaecologists

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or their carer or guardian.

Local commissioners and/or providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN:

Contents

Development of the guideline 5
Remit 5
What this guideline covers 5
Groups that will be covered 5
Key areas that will be covered in this update 5
Proposed outline for the guideline 6
What this guideline does not cover 8
Areas not covered by the guideline 8
Methods 9
Developing the review questions and outcomes 9
Reviewing research evidence 28
Type of studies and inclusion/exclusion criteria 28
Type of studies and inclusion/exclusion criteria 29
Methods of combining evidence 30
Appraising the quality of evidence 32
Intervention studies 32
Diagnostic test accuracy reviews 37
Qualitative reviews 38
Evidence statements 38
Reviewing economic evidence 39
Inclusion and exclusion of economic studies 39
Appraising the applicability and quality of economic evidence 40
Health economic modelling 40
Cost effectiveness criteria 41
Developing recommendations 41
Guideline recommendations 41
Research recommendations 41
Validation process 41
Updating the guideline 42
Funding 42
References 42

Development of the guideline

Remit

The National Institute for Health and Care Excellence (NICE) commissioned the National Guideline Alliance (NGA) to update the guideline on urinary incontinence in women: management (CG171).

The remit for this guideline update is to revise the NICE guideline on the urinary incontinence in women and expand the guideline to include pelvic organ prolapse.

What this guideline covers

Groups that will be covered

- Women (aged 18 and over) with urinary incontinence.
- Women (aged 18 and over) with pelvic organ prolapse. (To be included in the update but not covered in the existing guideline.)
- Women (aged 18 and over) with complications associated with insertion of mesh for treating stress urinary incontinence or pelvic organ prolapse. (To be included in the update but not covered in the existing guideline.)

Specific consideration will be given to:

- older women
- women with physical disabilities
- women with cognitive impairment
- women considering future pregnancy.

Key areas that will be covered in this update

We will look at evidence in the areas below when developing this update. We will consider making new recommendations or updating existing recommendations in these areas only.

- Assessing stress urinary incontinence: urodynamic testing.
- Alternative conservative management options for urinary incontinence: absorbent products.
- Drugs for overactive bladder.
- Invasive procedures for overactive bladder.
- Surgical procedures for stress urinary incontinence.
- Multidisciplinary team.
- Assessing pelvic organ prolapse.
- Managing pelvic organ prolapse.
- Managing coexisting urinary incontinence and pelvic organ prolapse.
- Assessing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse.

- Managing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse.

Note that guideline recommendations for medicines will normally fall within licensed indications; exceptionally, and only if clearly supported by evidence, use outside a licensed indication may be recommended. The guideline will assume that prescribers will use a medicine's summary of product characteristics to inform decisions made with individual patients.

Proposed outline for the guideline

Table 1 below outlines all the areas that will be included in the guideline. It sets out what NICE plans to do for each area in this update.

Table 1: Outline of areas included in the guideline

Area in the guideline	What NICE plans to do
Assessment and investigation of UI: • history taking and physical examination • pelvic floor muscle assessment • urine testing • assessment of residual urine • referral • symptom scoring and quality-of-life assessment • bladder diaries • pad testing • other tests of urethral competence • cystoscopy • imaging	No evidence review: retain recommendations from existing guideline
Assessment and investigation of UI: information provision	No evidence review: no recommendations in existing guideline owing to lack of evidence
Assessment and investigation of UI: urodynamic testing	Review evidence: update existing recommendations as needed
Conservative management of UI: • lifestyle interventions • physical therapies • behavioural therapies • neurostimulation • alternative conservative management options – urinals and toileting aids – catheters	No evidence review: retain recommendations from existing guideline

– products to prevent leakage – complementary therapies – preventive use of conservative therapies • progress of treatment	
Conservative management of UI: • alternative conservative management options: pessaries • optimal sequence and timescales for conservative therapies	No evidence review: no recommendations in existing guideline owing to lack of evidence
Conservative management of UI: • alternative conservative management options for UI: absorbent products	Review evidence: update existing recommendations as needed
Pharmacological treatment for UI: desmopressin, duloxetine and oestrogens	No evidence review: retain recommendations from existing guideline
Pharmacological treatment for UI: diuretics	No evidence review: no recommendations in existing guideline owing to lack of evidence
Pharmacological treatment for UI: drugs for OAB	Review evidence: update existing recommendations as needed
Invasive procedures for OAB: • percutaneous sacral nerve stimulation • augmentation cystoplasty • urinary diversion	No evidence review: retain recommendations from existing guideline
Invasive procedures for OAB: • detrusor myectomy • vanilloid receptor agonists • sequence of surgical procedures for overactive bladder – economic evaluation	No evidence review: no recommendations in existing guideline owing to lack of evidence
Invasive procedures for OAB: botulinum toxin	Review evidence: update existing recommendations as needed
Surgical procedures for stress UI	Review evidence: update existing recommendations as needed
Multidisciplinary team	Review evidence: update existing recommendations as needed
Competence of surgeons performing operative procedures for UI in women	Remove: refer to professional body competence standards

Assessment and investigation of POP: assessment	Review evidence: update existing recommendations as needed
Conservative management of POP: • lifestyle interventions • other conservative management options	Review evidence: new area in the guideline
Pharmacological treatment for POP	Review evidence: new area in the guideline
Surgical procedures for POP	Review evidence: new area in the guideline
Managing coexisting UI and POP	Review evidence: new area in the guideline
Assessing complications associated with mesh surgery for stress UI or POP	Review evidence: new area in the guideline
Managing complications associated with mesh surgery for stress UI or POP	Review evidence: new area in the guideline

MDT, multidisciplinary team; OAB, overactive bladder; POP, pelvic organ prolapse; UI, urinary incontinence.

Recommendations in areas that are being retained from the existing guideline may be edited to ensure that they meet current editorial standards, and reflect the current policy and practice context.

What this guideline does not cover

Areas not covered by the guideline

- Information provision and consent for women considering surgical intervention for stress urinary incontinence or pelvic organ prolapse – this is being specifically addressed in reviews by NHS England and NHS Scotland.
- Incontinence associated with neurological disease.
- Rectal prolapse.
- Fistula, except in relation to complications associated with mesh surgery.
- Women who had surgical management of congenital anomalies of the lower genitourinary tract as children.
- Faecal incontinence.
- Urinary incontinence associated with pregnancy.
- Causes of and risk factors for pelvic organ prolapse.
- Causes of and risk factors for postoperative incontinence after prolapse surgery.
- Assessing complications after non-mesh surgery for urinary incontinence and pelvic organ prolapse.
- Managing complications after non-mesh surgery for urinary incontinence and pelvic organ prolapse.
- Managing complications after mesh surgery that are not caused by mesh surgery.

Methods

This chapter sets out in detail the methods used to review the evidence and to generate recommendations in the guideline. This guideline was developed using the methods described in the 2014 NICE guidelines manual.

Declarations of interest were recorded according to the 2014 NICE conflicts of interest policy until 31st March 2018. From 1st April 2018 declarations of interest were recording according to the 2018 NICE conflicts of interest policy on declaring and managing interests for NICE advisory committees. Those interests declared until April 2018 were reclassified according to NICE's 2018 conflicts of interest policy (see Register of Interests).

For information on methods used to develop the evidence reviews not addressed in this guideline update, see guideline development methodology section in the 2013 guideline.

Developing the review questions and outcomes

The 22 review questions developed for this guideline were based on the key areas identified in the guideline scope. They were drafted by the NGA, and refined and validated by the guideline committee. They covered all areas of the scope and were signed-off by NICE. These questions are outlined in Table 2.

The review questions were based on the following frameworks:

- intervention reviews: population, intervention, comparator and outcome (PICO)
- diagnostic test accuracy reviews: population, index test, reference standard and outcome (PIRO)
- qualitative reviews: Population or problem, interest (i.e. defined event, activity, experience or process) and context (PICo)

These frameworks guided the development of the review protocols, the literature searching process, the critical appraisal and synthesis of evidence and facilitated the development of recommendations by the committee.

Full literature searches, critical appraisals and evidence reviews were completed for all review questions.

Description of review questions

Table 2: Description of review questions

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
1 Assessing stress urinary incontinence	A	Intervention	1.1 What is the value of urodynamic assessment in addition to clinical assessment before primary surgery for	**Critical outcomes:** 1. Continence status (improvement e.g. number of incontinent episodes per day in first 3 months after treatment)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
			stress urinary incontinence?	2. Adverse effects of urodynamic testing • urinary infection • dysuria • haematuria 3. Continence specific health-related quality of life (ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI and KHQ (all from previous guideline) and E-PAQ (new)) **Important outcomes:** 4. Adverse effects of SUI surgery • Urgency, urgency incontinence, voiding difficulties 5. Satisfaction • Patient Global Impression of Improvement 6. Change of management
2 Alternative conservative management options for urinary incontinence	B	Intervention	2.1 How often should alternative treatment options be reviewed for women who are using absorbent containment products?	**Critical outcomes:** 1. Skin breakdown, ulcers 2. Other procedures offered (i.e. surgery)/Women moving to an alternative treatment option 3. Incontinence specific health-related quality of life (e.g. ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI and KHQ and E-PAQ. **Important outcomes:** 4. Infection 5. Patient satisfaction
3 Drugs for overactive bladder	C	Intervention	3.1 What are the risks to cognitive function for women taking anticholinergic drugs for overactive bladder?	**Critical outcomes:** 1. Long-term cognitive impairment measured using validated tools only, including: • Abbreviated metal test score • General practitioner assessment of cognition

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• Mini-cog • Addenbrookes cognitive examination III • Montreal cognitive assessment • Mini mental state examination • 6-item cognitive impairment test 2. Falls **Important outcomes:** 3. Delirium 4. All-cause mortality
4 Invasive procedures for overactive bladder	D	Intervention	4.1 What is the value of urodynamic assessment before botulinum toxin type A treatment?	**Critical outcomes:** 1. Continence status (improvement e.g. number of incontinent episodes per day in first 3 months after treatment) 2. Adverse effects of urodynamic testing • urinary infection • dysuria • haematuria 3. Continence specific health-related quality of life (ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI and KHQ (all from previous guideline) and E-PAQ (new)) **Important outcomes:** 4. Adverse effects of SUI surgery • Urgency, urgency incontinence, voiding difficulties 5. Adverse effects of Botulinum toxin (UTI, requirement for self-catheterisation) 6. Satisfaction • Patient global impression of improvement 7. Change of management

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
4 Invasive procedures for overactive bladder	D	Intervention	4.2 What is the most effective dose of botulinum toxin type A for treating overactive bladder?	**Critical outcomes:** 1. Continence status (e.g. number of incontinent episodes per day in first 3 months after treatment) 2. Continence specific Quality of life (ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI and KHQ (all from previous guideline) and E-PAQ (new)) 3. Requirement for self-catheterisation or indwelling catheterisation **Important outcomes:** 1. Symptom reduction (e.g. number of urgency and frequency episodes per day in first 3 months after treatment) 2. Adverse effects (e.g. urinary infection, retention) 3. Satisfaction (patient rated improvement)
5 Surgical procedures for stress urinary incontinence	E	Intervention	5.1 What is the most effective surgical management of stress urinary incontinence, including mesh and non-mesh procedures?	**Critical outcomes:** 1. Continence specific health-related quality of life • ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI and KHQ, E-PAQ • Sexual function (PISQ-12) 2. Adverse events (immediate post-operative or perioperative) • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel) 3. Complications • Pain • Mesh erosion or extrusion (vaginal, bladder, urethra)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• Fistula • Need for catheterisation (include voiding dysfunction, e.g. retention, slow stream, incomplete emptying) • Infection (recurrent UTI, wound) • De novo overactive bladder symptoms (clinically-established but possibly confirmed by urodynamics) • Urge incontinence • Frequency • Urgency • Nocturne • Occurrence of POP • Wound complications (hernia) Complications will be stratified as follows: • Short-term: complications occurring up to 1 year (i.e., ≤ 1 year); • Medium-term: complications occurring after 1 year, and up to 5 years (i.e., >1 to ≤ 5 years); and • Long-term: complications occurring after 5 years (i.e., > 5 years) **Important outcomes:** 4. Change in continence status • Subjective report • Objective cure rate • Negative stress (cough) test

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• Number of incontinence episodes per day 5. Patient satisfaction, patient reported improvement • Patient global impression of improvement 6. Repeat surgery (for UI or POP, or mesh complications)
5 Surgical procedures for stress urinary incontinence	E	Intervention	5.2 What is the effectiveness of surgical management of stress urinary incontinence (including mesh and non-mesh procedures), compared to pelvic floor muscle training?	**Critical outcomes:** 1. Continence-specific health-related quality of life • Specific scales: ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI, KHQ, and E-PAQ (new) (Total scores if available) • Sexual function: PISQ 2. Change in continence status • Subjective report • Objective cure rate • Negative stress (cough) test • Number of incontinence episodes per day 3. Patient satisfaction, patient reported improvement • Patient global impression of improvement • Number of women who are satisfied **Important outcomes:** 4. Adverse events (immediate post-op or perioperative) • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				5. Long-term complications (>12 months) • Pain • Mesh erosion or extrusion (vaginal, bladder, urethra) • Fistula • Need for catheterisation • Infection (recurrent UTI, wound) • De novo overactive bladder symptoms • Occurrence of POP • Wound complications (hernia) 6. Repeat surgery (for UI or POP, or mesh complications)
6 Multidisciplinary team	F	Intervention	6.1 What is the most effective composition of a multidisciplinary team for the assessment and management of simple and complex cases including mesh complications?	**Critical outcomes:** 1. Change in management decisions 2. Health-related quality of life (specific to UI or POP). **Important outcomes:** 3. Patient satisfaction
7 Assessing pelvic organ prolapse	G	Diagnostic	7.1 What is the most effective strategy for assessing pelvic organ prolapse?	**Critical Outcomes** 1. Sensitivity 2. Specificity 3. Positive likelihood ratio 4. Negative likelihood ratio **Important outcomes:** 5. Patient satisfaction 6. Symptom improvement • Self-reported • Assessed using validated questionnaire 7. Change in management option? 8. Pain associated with test/assessment 9. Anxiety associated with test/assessment

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
8 Managing pelvic organ prolapse	H	Intervention	8.1 What lifestyle interventions are effective for managing pelvic organ prolapse?	**Critical outcomes:** 1. Improvement in symptoms • Self-reported symptoms • Questionnaires: o POP-SS o ICIQ-VS o EPAQ o PFIQ-7/PFDI-20 2. Patient satisfaction (measured by PFDI, or patient reported) 3. Health-related quality of life (measured by EQ-5D) **Important Outcomes** 4. Sexual function (PISQ) 5. Adverse events 6. Anatomical assessment of POP (assessed by POP-Q)
8 Managing pelvic organ prolapse	H	Intervention	8.2 What is the effectiveness of topical oestrogen for managing pelvic organ prolapse with vaginal atrophy?	**Critical outcomes:** 1. Improvement in symptoms: • Self-reported symptoms • Questionnaires: o POP-SS o ICIQ-VS o EPAQ o PFIQ-7/PFDI-20 2. Patient satisfaction (measured by PFDI, patient reported) 3. Health-related quality of life (measured by EQ-5D) **Important outcomes:** 4. Sexual function (PIS-Q) 5. Adverse events (post-menopausal bleeding, breast symptoms pain/tenderness, pelvic discomfort and pain, discharge, allergic reaction) 6. Anatomical assessment of POP (assessed by POP-Q)
8 Managing pelvic organ prolapse	H	Intervention	8.3 What are the most effective conservative management options (for example, pelvic floor exercises and pessaries) for pelvic organ prolapse?	**Critical outcomes:** 1. Improvement in symptoms • Self-reported symptoms • Questionnaires: POP-SS, EPAQ, PFDI-20 2. Patient satisfaction (measured by PFDI, patient reported)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				3. Health related quality of life (measured by EQ-5D, ICIQ-VS, PFIQ-7) **Important outcomes:** 4. Sexual function (PIS-Q) 5. Adverse events 6. Anatomical assessment of POP (assessed by POP-Q)
8 Managing pelvic organ prolapse	I	Intervention	8.4 What are the most effective surgical management options (including mesh and non-mesh procedures) for pelvic organ prolapse?	**Critical outcomes:** 1. Health related quality of life (measured through validated scales only) 2. Adverse events • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel) 3. Complications • Pain • Mesh erosion or extrusion (bladder, vagina, bowel, urethra) • Fistula • Bladder function ○ Stress UI ○ Urge incontinence ○ Voiding difficulty • Bowel function ○ Faecal incontinence ○ Obstructed defecation ○ Constipation • Sexual function ○ De novo dyspareunia ○ Apareunia ○ Prolapse and incontinence sexual questionnaire • Recurrence of any POP ○ Same compartment

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				○ Different compartment Complications will be stratified as follows: • Short-term: complications occurring up to 1 year (i.e., ≤ 1 year); • Medium-term: complications occurring after 1 year, and up to 5 years (i.e., > 1 year and ≤ 5 years); and • Long-term: complications occurring after 5 years (i.e., > 5 years) **Important outcomes:** 4. Cure/Prolapse • Subjective report or affirmation • Objective examination (POP-Q staging) 5. Patient satisfaction 6. Repeat surgery (for UI or POP, mesh complications)
8 Managing pelvic organ prolapse	I	Intervention	8.5 What is the role of surgery to prevent postoperative urinary incontinence in women having surgery for pelvic organ prolapse, including the sequence of interventions?	**Critical outcomes:** 1. Change in continence status • Self-reported symptoms • Objective cure rate • Negative stress (cough) test • Number of incontinence episodes per day 2. Long-term complications (> 12 months) • Pain • Mesh erosion or extrusion (vaginal, bladder, urethra) • Fistula • Need for catheterisation • Infection (recurrent UTI, wound)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• De novo overactive bladder symptoms • Occurrence of POP • Wound complications (hernia) 3. Repeated surgery for UI, POP or mesh complications **Important outcomes:** 4. Continence specific health-related quality of life (ICIQ, BFLUTS, I-QOL, SUIQQ, UISS, SEAPI-QMM, ISI, KHQ and E-PAQ) 5. Adverse events (immediate post-op or perioperative) • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel) 6. Patient satisfaction • Patient reported improvement • Patient global impression of improvement
8 Managing pelvic organ prolapse	I	Intervention	8.6 What is the effectiveness of surgical options for pelvic organ prolapse, compared to pessaries?	**Critical outcomes:** 1. Health related quality of life (measured through validated scales only) 2. Adverse events • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel) 3. Long-term adverse events • Pain • Mesh erosion or extrusion (bladder, vagina, bowel, urethra) • Fistula • Bladder function o Stress UI

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				○ Urge incontinence ○ Voiding difficulty • Bowel function ○ Faecal incontinence ○ Obstructed defecation ○ Constipation • Sexual function ○ De novo dyspareunia ○ Apareunia ○ Prolapse and incontinence sexual questionnaire • Recurrence of any POP ○ Same compartment ○ Different compartment **Important outcomes:** 4. Cure/Prolapse • Subjective report or affirmation • Objective examination (POP-Q staging) 5. Patient satisfaction 6. Need for subsequent surgery (for UI or POP, mesh complications)
9 Managing coexisting urinary incontinence and pelvic organ prolapse	J	Intervention	9.1 What is the most effective surgical management for women with both stress urinary incontinence and pelvic organ prolapse, including the sequence of interventions?	**Critical outcomes:** 1. Change in continence status • Self-reported symptoms • Objective cure rate (to be examined in NMA and pairwise results to be presented there) • Negative stress (cough) test • Pad test (1-hr or 24-hr)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• Number of incontinence episodes per day 2. Repeat surgery (for UI or POP, or mesh complications) 3. Long-term complications (>12 months) • Pain • Mesh erosion or extrusion (vaginal, bladder, urethra) • Fistula • Need for catheterisation • Infection (recurrent UTI, wound) • De novo overactive bladder symptoms • Occurrence of POP • Wound complications (hernia) **Important outcomes:** 4. Adverse events (immediate post-op or perioperative) • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel) 5. Incontinence specific health-related quality of life • Sexual function • King's Health Questionnaire 6. Patient satisfaction, patient reported improvement • Patient global impression of improvement
10 Assessing complications associated with mesh	K	Intervention	10.1 What is the most effective strategy for assessing	**Critical outcomes:** 1. Patient satisfaction • PGI-I • PGI-S

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
surgery for stress urinary incontinence or pelvic organ prolapse			complications (for example, vaginal complications, sexual dysfunction, pain, urinary symptoms and bowel symptoms) after mesh surgery?	• Self-reported 2. Symptoms and Quality of life • Self-reported symptoms (all complications) • ePAQ (all complications) • PIS-Q (sexual dysfunction) • For UI ○ ICIQ ○ BFLUTS ○ I-QOL ○ SUIQQ ○ UISS ○ SEAPI-QMM ○ ISI ○ KHQ • For POP ○ POP-SS ○ ICIQ-VS ○ PFIQ-7/PFDI-20 3. Pain relief (measured using validated scales specific to UI and/or POP; in their absence, we will consider the use of VAS or the number of women experiencing – or not-improvement of their pain (i.e., a dichotomous outcome) 4. Adverse events associated with testing **Important outcomes:** 5. Change in clinical management
11 Managing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse	L	Intervention	11.1 What are the most effective management options for vaginal complications (including exposure, extrusion and infection) after mesh surgery?	**Critical outcomes:** 1.Continued or repeated exposure/extrusion/infection 2.Adverse events (immediate post-op or perioperative): • Severe bleeding requiring a blood transfusion • Internal organ injury (to bladder or bowel)

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				3.Long-term complications (> 12 months): • Pain • Mesh erosion or extrusion • Fistula • Need for catheterisation • Infection • De novo overactive bladder symptoms • Sexual dysfunction • Wound complications (infection and tissue breakdown) **Important outcomes:** 4. Health-related quality of life (validated scales only) 5. Patient satisfaction • Patient reported improvement • Patient Global Impression of Improvement 6. Repeat surgery (for mesh complications) 7. Recurrence of urinary incontinence or prolapse
11 Managing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse	L	Intervention	11.2 What are the most effective management options for sexual dysfunction after mesh surgery?	**Critical outcomes:** 1. Sexual function (measured using validated scales such as PISQ or ePAQ) 2. Adverse events (immediate post-op or perioperative): • Severe bleeding requiring blood transfusion • Unintentional internal organ injury 3. Patient satisfaction • Patient reported improvement • Patient Global Impression of Improvement

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				Important outcomes: 4. Health-related quality of life 5. Repeat surgery (for UI or POP, or mesh complications) 6. Long-term complications (> 12 months): • Pain • Fistula • Infection • Wound complications 7. Partner satisfaction
11 Managing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse	L	Intervention	11.3 What are the most effective management options for pain after mesh surgery?	**Critical outcomes:** 1. Pain (measured through a validated scale; appropriate MIDs to use if available will be identified through consultation with the GC) 2. Patient satisfaction • Patient-reported improvement • Patient Global Impression of Improvement 3. Adverse events (immediate post-op or perioperative): • Severe bleeding requiring blood transfusion • Unintentional internal organ injury **Important outcomes:** 4. Health-related quality of life 5. Repeat surgery (for UI or POP, or mesh complications) 6. Long-term complications (> 12 months) • Pain • Fistula • Infection • Would complications • Mesh erosion or extrusion

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• De novo overactive bladder symptoms • Sexual dysfunction • Need for catheterisation 7. Recurrence of urinary incontinence or prolapse
11 Managing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse	L	Intervention	11.4 What are the most effective management options for urinary complications after mesh surgery?	**Critical outcomes:** 1. Continued or repeated urinary complications (as per above including mesh) 2. Adverse events (immediate post-op or perioperative): • Severe bleeding requiring a blood transfusion • Unintentional Internal organ injury (bladder or bowel or ureter) 3.Long-term complications (> 12 months): • Pain • Fistula • Need for catheterisation • Infection • De novo overactive bladder symptoms • Wound complications • Urinary incontinence **Important outcomes:** 4. Continence specific health-related quality of life: • ICIQ • BFLUTS • I-QOL • SUIQQ • UISS • SEAPI-QMM, • ISI • KHQ • E-PAQ 5. Patient satisfaction

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				• Patient reported improvement • Patient Global Impression of Improvement 6. Repeat surgery (for UI or POP, or mesh complications)
11 Managing complications associated with mesh surgery for stress urinary incontinence or pelvic organ prolapse	L	Intervention	11.5 What are the most effective management options for bowel symptoms after mesh surgery?	**Critical outcomes:** 1. Reduction in bowel symptoms 2. Adverse events (immediate post-operative or peri-operative: • Severe bleeding requiring blood transfusion • Unintentional internal organ injury 3. Health-related quality of life **Important outcomes:** 4. Complications (more than 12 months): • Pain • Fistula • Infection • Wound complications • Mesh erosion or extrusion • Sexual dysfunction 5. Patient satisfaction 6. Repeat surgery for UI, POP or mesh complications 7. Recurrence of urinary incontinence or prolapse Complications will be stratified as follows: • Short-term: complications occurring after one year or less (≤ 1 year) • Medium-term: complications occurring after one year and up to five

Chapter or section from the scope	Location in Evidence Reports	Type of review	Review question	Outcomes
				years (> 1 year and ≤ 5 years) Long-term: complications occurring after 5 years (> 5 years)

BFLUTS: Bristol Female Lower Urinary Tract Symptoms; E-PAQ: Electronic, Personal Assessment Questionnaire; ICIQ: International Consultation on Incontinence Questionnaire; I -QOL: Urinary Incontinence-Quality of Life Questionnaire; ICIQ-VS: International Consultation on Incontinence Questionnaire vaginal symptoms; ISI: Incontinence Symptom Index; KHQ: King's Health Questionnaire; NMA: Network Meta-Analysis; PFDI-20: Pelvic Floor Disability Index; PFIQ-7: Pelvic Floor Impact Questionnaire; PGI-I: PGI-I: Patients Global Impression of Improvement; PGI-S: Patients Global Impression of Severity; PISQ-12: Pelvic Organ Prolapse/incontinence Sexual Questionnaire; POP: Pelvic Organ Prolapse; POP-Q: Pelvic Organ Prolapse Quantification System; POP-SS: Pelvic Organ Prolapse Symptom Score; SEAPI-QMM: stress-related leak, emptying, anatomy, protection, inhibition, quality of life, mobility and mental status incontinence classification system; SUI: Stress Urinary Incontinence; SUIQQ: Stress and Urgency Incontinence and Quality of Life Questionnaire; UI: Urinary Incontinence; UISS: Urinary Incontinence Severity Score; UTI: Urinary Tract Infection; VAS: Visual Analogue Score.

Searching for evidence

Clinical search literature

Systematic literature searches were undertaken to identify all published clinical evidence relevant to the review questions.

Databases were searched using relevant medical subject headings, free-text terms and study type filters where appropriate. Studies published in languages other than English were not reviewed. All searches were conducted in MEDLINE, Embase and The Cochrane Library, with some additional database searching in AMED, PsycINFO and CINAHL for certain topic areas. The literature search strategies can be found in appendix B in each Evidence Report.

Searches were initially undertaken between March 2017 and March 2018 and re-runs performed in June 2018 were prioritised for the two surgical intervention topics (evidence reports E and I). These two topics were prioritised as a result of their importance to the guideline and due to the public concern with mesh procedures.

Search strategies were quality assured by cross-checking reference lists of highly relevant papers, analysing search strategies in other systematic reviews and asking the group members to highlight any additional studies. The questions, the study types applied and the databases searched can be found in appendix B in each Evidence Report. The years covered can be found in the review protocols.

Searches for grey literature or unpublished literature were not routinely undertaken, however some grey literature searching was undertaken for the Multidisciplinary

Teams (MDT), service delivery topic. Searches for electronic, ahead-of-print publications were not routinely undertaken.

During the scoping stage, a search was conducted for guidelines, systematic reviews and reports on websites of organisations relevant to the topic. All references suggested by stakeholders at the scoping consultation were considered to determine whether they met the inclusion criteria of the reviews.

Health economics search literature

A global search of economic evidence was undertaken in Medline, Embase, HTA database and NHS EED in November 2016 and re-run in June 2018. Evidence resulting from the search was screened to reflect the final dates of the searches that were undertaken for the clinical reviews (see review protocols).

Further to the database searches, the committee was contacted with a request for details of relevant published and unpublished studies of which they may have had knowledge; reference lists of key identified studies were also reviewed for any potentially relevant studies.

The search strategy for existing economic evaluations combined terms capturing the target condition (UI/POP) and, for searches undertaken in MEDLINE and EMBASE, terms capturing UI/POP and economic evaluations. No restrictions on language or setting were applied to any of the searches, but a standard exclusions filter was applied (letters, animals, etc.). Full details of the search strategies are presented in appendix B of each Evidence Report and in Supplementary Material D – Health Economic Literature.

Call for evidence

No call for evidence was made.

Reviewing research evidence

Type of studies and inclusion/exclusion criteria

The evidence was reviewed following these steps.

- Potentially relevant studies were identified for each review question from the relevant search results by reviewing titles and abstracts. Full papers were then obtained.
- Full papers were reviewed against pre-specified inclusion and exclusion criteria as outlined in the review protocols (in appendix A of each evidence review chapter).
- Key information was extracted on the study's methods, according to the factors specified in the protocols and results. These were presented in summary tables (in each review chapter) and evidence tables (in appendix D of each evidence review chapter).
- Relevant studies were critically appraised using the appropriate checklist as specified in Developing NICE guidelines: the manual 2014

- Summaries of evidence were generated by outcome (included in the relevant review chapters) and were presented to the committee as follows.
  - Randomised and non-randomised comparative studies: meta-analysis was carried out where appropriate and results were reported in Grading of Recommendations Assessment, Development and Evaluation (GRADE) profiles (for intervention reviews).
  - Non-comparative observational studies: data regarding medium- and long-term complications of surgical interventions for pelvic organ prolapse, and data on long-term complications of surgical interventions for stress urinary incontinence, were combined and presented in summary tables as weighted averages. Individual studies were assessed for risk of bias using the appropriate study checklist.
  - Qualitative studies: each study was summarised by theme and themes were then presented in summary tables with quality ratings based on the study checklists.
- All drafts of reviews were checked by a senior reviewer.

Type of studies and inclusion/exclusion criteria

For intervention reviews in this guideline, randomised controlled trials (RCT) were prioritised because they are considered the most robust type of study design for unbiased estimate of intervention effects. Non-randomised comparative observational studies (e.g. cohort) were considered if there was no or very little RCT evidence. Non-comparative studies were considered in the reviews of surgical interventions for SUI and/or POP to estimate the medium- and/or long-term rates of specific complications (e.g. pain).

In the qualitative reviews, studies using focus groups, or structured or semi-structured interviews were considered for inclusion. Survey data or other types of questionnaires were only included if they provided analysis from open-ended questions, but not if they reported descriptive quantitative data only.

For quality assurance of study identification, titles and abstracts of identified studies were screened by two reviewers for inclusion against criteria, until a good inter-rater reliability was observed (percentage agreement =>90% or Kappa statistics, K>0.60). Initially 10% of references were double-screened. If inter-rater agreement was good then the remaining references were screened by one reviewer. All primary-level studies included after the first scan of citations were acquired in full and re-evaluated for eligibility at the time they were entered into a study database (standardised template created in Microsoft Excel). At least 10% of data extraction were double-coded. Discrepancies or difficulties with coding were resolved through discussion between reviewers or the opinion of a third reviewer was sought. Non-English-language papers were excluded (unless data were obtained from an existing review). For further details, please refer to Appendix A of the relevant Evidence Report.

Methods of combining evidence

Data synthesis for intervention reviews

Pairwise meta-analysis

Pairwise meta-analysis of homogenous randomised trails was done using Review Manager 5 (RevMan 5) software. For binary outcomes, such as occurrence of adverse events, the Mantel-Haenszel method of statistical analysis was used to calculate risk ratios (relative risks, RR) with 95% confidence intervals (CI).

For continuous outcomes, measures of central tendency (mean) and variation (standard deviation (SD)) are required for meta-analysis. Data for continuous outcomes (such as health-related quality of life score or length of hospital stay) were analysed using an inverse-variance method for pooling weighted mean differences.

Statistical heterogeneity was assessed by visually examining the forest plots, and by considering the chi-squared test for significance with heterogeneity defined as a $p<0.1$ or an I-squared inconsistency statistic value of 50% or more. Where heterogeneity was present, predefined subgroup analyses were performed. If the heterogeneity still remained, a random effects (DerSimonian 2015) model was employed to provide a more conservative estimate of the effect.

Results from multiple observational studies of the same comparison were not pooled but presented as a range of effects. This was due the high risk of selection bias in observational studies whereby differences in participant characteristics between treatment arms leads to a biased estimate of treatment effect.

Forest plots were generated to present the results for outcomes with more than one study (please see appendix E of each intervention evidence review).

In the evidence reviews on surgical interventions for SUI and POP (evidence reports E and I), and RCT data were not available for all complications post 12 months; therefore, for the long-term complications of SUI surgery, and for medium- and long-term complications following POP surgery, data were extracted from a variety of study types (RCT, cohort studies and/or case series). The data were extracted as number of events for each complication, and the weighted average (weighted by sample size) calculated. For the data on complications following POP surgery weighted averages were grouped according to placement of mesh (i.e. abdominal and vaginal mesh surgery). Data on SUI surgery was grouped according to type of intervention.

Network meta-analysis

In the evidence review looking at the effectiveness of surgical management options (including mesh and non-mesh procedures) for anterior pelvic organ prolapse recurrence at the same site outcome, the evidence synthesis used network meta-analytic techniques with the network meta-analysis (NMA) review protocol presented in the relevant chapter I, appendix N.

As is the case for ordinary pairwise meta-analysis, NMA may be conducted using either fixed or random effect models. A fixed effect model typically assumes that there is no variation in relative effects across trials for a particular pairwise comparison and any observed differences are solely due to chance. For a random

effects model, it is assumed that the relative effects are different in each trial but that they are from a single common distribution. The variance reflecting heterogeneity is often assumed to be constant across trials.

In a Bayesian analysis, for each parameter the evidence distribution is weighted by a distribution of prior beliefs. The Markov Chain Monte Carlo (MCMC) algorithm was used to generate a sequence of samples from a joint posterior distribution of 2 or more random variables and is particularly well adapted to sampling the treatment effects (known as a posterior distribution) of a Bayesian network. A prior distribution was used to maximise the weighting given to the data and to generate the posterior distribution of the results.

For the analyses, a series of burn-in simulations were run to allow the posterior distributions to convergence and then a further simulations were run to produce the posterior outputs. Convergence was assessed by examining the history, autocorrelation and Brooks-Gelman-Rubin plots.

Goodness-of-fit of the model was also estimated by using the posterior mean of the sum of the deviance contributions for each item by calculating the residual deviance and deviance information criteria (DIC). If the residual deviance was close to the number of unconstrained data points (the number of trial arms in the analysis) then the model was explaining the data at a satisfactory level. The choice of a fixed effect or random effects model can be made by comparing their goodness-of-fit to the data.

The consistency between direct and indirect evidence can be assessed in closed treatment loops within the network. These closed treatment loops are regions within a network where direct evidence is available on at least 3 different treatments that form a closed ‘circuit’ of treatment comparisons (for example, A versus B, B versus C, C versus A). If closed treatment loops existed then discrepancies between direct and indirect evidence was assessed. The consistency checks were undertaken by TSU, University of Bristol and are summarised in the relevant chapter, appendix S.

Treatment specific posterior effects were generated for every possible pair of comparisons by combining direct and indirect evidence in each network. The probability that each treatment is best, based on the proportion of Markov chain iterations in which the treatment effect for an intervention is ranked best, second best and so forth. This was calculated by taking the treatment effect of each intervention compared to the reference treatment and counting the proportion of simulations of the Markov chain in which each intervention had the highest treatment effect.

One of the main advantages of the Bayesian approach is that the method leads to a decision framework that supports decision making. The Bayesian approach also allows the probability that each intervention is best for achieving a particular outcome, as well as its ranking, to be calculated.

We adapted standard fixed and random effects Binomial models with cloglog link available from NICE Decision Support Unit (DSU) technical support document number 2: http://nicedsu.org.uk/wp-content/uploads/2017/05/TSD2-General-meta-analysis-corrected-2Sep2016v2.pdf

For further description of the model used, specific methods, outcomes and the results of the NMA please see chapter I.

The quality assurance of all the NMA work was undertaken by TSU, University of Bristol.

The guideline committee also considered the published NMA (Brazzelli 2018 – in preparation) that examined the effectiveness of surgical options for stress urinary incontinence. The version of Brazzelli (2018) that was considered by the NICE guideline committee was a draft version of the manuscript dated July 2018. That version is yet to complete the editorial review process in line with the National Institute for Health Research (NIHR) Journals Library policy. This project was funded by the Health Technology Assessment (HTA 15/09/06) and will be published in full in the *Health Technology Assessment* journal. Further information available at: https://www.journalslibrary.nihr.ac.uk/programmes/hta/150906/#/

Brazzelli (2018) presents independent research commissioned by the NIHR. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, MRC, CCF, NETSCC, the Programme Grants for Applied Research programme or the Department of Health.

Data synthesis for diagnostic test accuracy reviews

Meta-analysis of diagnostic test accuracy was conducted using either single or multiple test analysis, sensitivity and specificity plots were generated to present the results, (please see appendix E of each diagnostic test accuracy evidence review chapter).

Appraising the quality of evidence

Intervention studies

GRADE methodology

For intervention reviews, the evidence for outcomes from the included RCTs was evaluated and presented using GRADE, which was developed by the international GRADE working group.

The software developed by the GRADE working group (GRADEpro) was used to assess the quality of each outcome, taking into account individual study quality factors and the meta-analysis results. The clinical/economic evidence profile tables include details of the quality assessment and pooled outcome data, where appropriate, an absolute measure of intervention effect and the summary of quality of evidence for that outcome. In this table, the columns for intervention and control indicate summary measures of effect and measures of dispersion (such as mean and SD or median and range) for continuous outcomes and frequency of events (n/N; the sum across studies of the number of patients with events divided by sum of the number of completers) for binary outcomes. Reporting or publication bias was only taken into consideration in the quality assessment and included in the clinical evidence profile tables if it was apparent.

The selection of outcomes for each review question was decided when each review protocol was discussed with the guideline committee, and was informed by committee discussion and key papers.

The evidence for each outcome in the intervention reviews was examined separately for the quality elements listed and defined in Table 3. Each element was graded using the quality levels listed in Table 4.

The main criteria considered in the rating of these elements are discussed below. Footnotes were used to describe reasons for grading a quality element as having serious or very serious limitations. The ratings for each component were summed to obtain an overall assessment for each outcome (Table 5).

Table 3: Description of quality elements in GRADE for intervention reviews

Quality element	Description
Risk of bias (study limitations)	Limitations in the study design and implementation may bias the estimates of the treatment effect. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of results or findings.
Indirectness	Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, or recommendation made, such that the effect estimate is changed. This is also related to applicability or generalisability of findings.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold.
Publication bias	Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.

Table 4: Levels of quality elements in GRADE

Levels of quality elements in GRADE	Description
None/no serious	There are no serious issues with the evidence.
Serious	The issues are serious enough to downgrade the outcome evidence by 1 level.
Very serious	The issues are serious enough to downgrade the outcome evidence by 2 levels.

Table 5: Levels of overall quality of outcome evidence in GRADE

Overall quality of outcome evidence in GRADE	Description
High	Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low	Any estimate of effect is very uncertain.

Assessing risk of bias in intervention reviews

Bias is a systematic error, or a consistent deviation from the truth in the results. When a risk of bias is present the true effect can be either under- or over-estimated.

Risk of bias in RCT studies was assessed using the Cochrane Risk of Bias Tool (see appendix H in Developing NICE guidelines: the manual 2014

The different sources of bias in RCT studies in the Cochrane risk of bias tool fall into the following 5 categories: selection bias, performance bias, attrition bias, detection bias and reporting bias.

The Risk Of Bias In Non-randomised Studies – of Interventions (ROBINS-I) tool was used to assess risk of bias in other cohort and non-comparative studies (see appendix H in Developing NICE guidelines: the manual 2014

The different sources of bias in non-randomised studies in the ROBINS-I tool fall into the following 7 categories: confounding bias, selection bias, classification of interventions bias, deviations from intended interventions bias, missing data bias, measurement of outcomes bias, and selective reporting bias.

It should be noted that a study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether this poor design will impact on the estimation of the intervention effect.

For risk of bias, outcomes were downgraded if the randomisation and/or allocation concealment methods were unclear or inadequate. Outcomes were also downgraded if no attempts were made to blind the assessors or participants except in cases where blinding is not possible, impractical and/or unethical. Outcomes were also downgraded if there was considerable missing data (see below).
Handling missing data:

- where possible, an intention to treat approach was used
- outcomes were downgraded if there was a dropout of more than 20%, or if there was a difference of >20% between the groups.

Assessing inconsistency in intervention reviews

Inconsistency refers to unexplained heterogeneity of results of meta-analysis. When estimates of the treatment effect vary widely across studies (that is, there is heterogeneity or variability in results), this suggests true differences in underlying effects. Inconsistency is, thus, only applicable when statistical meta-analysis is conducted (that is, results from different studies are pooled). When outcomes were derived from a single study the rating ‘no serious inconsistency’ was used when assessing the domain, as per GRADE methodology (Santesso 2016).

Statistical heterogeneity was assessed by calculating the I-squared statistic for the meta-analysis. I-squared values of equal to or more than 50% and 80% were considered to indicate high and very high heterogeneity, respectively. When high or very high heterogeneity was observed, possible reasons for it were explored and subgroup analyses were performed as pre-specified in the review protocol.

The quality of the evidence was downgraded in GRADE by 1 (I-squared ≥ 50%) or 2 (I-squared ≥ 80%) levels for the domain of inconsistency, depending on the extent of heterogeneity in the results.

Assessing indirectness in intervention reviews

Directness refers to the extent to which the populations, intervention, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is important when these differences are expected to contribute to a difference in effect size, or may affect the balance of harms and benefits considered for an intervention.

Assessing imprecision and clinical significance in intervention reviews

Imprecision in guidelines concerns whether the uncertainty (CI) around the effect estimate means that it is not clear whether there is a clinically important difference between interventions or not (that is, whether the evidence would clearly support one recommendation or appear to be consistent with several different types of recommendations). Therefore, imprecision differs from the other aspects of evidence quality because it is not really concerned with whether the point estimate is accurate or correct (has internal or external validity). Instead, it is concerned with the uncertainty around the point estimate actually is. This uncertainty is reflected in the width of the CI.

The 95% CI is defined as the range of values within which the population mean value will fall on 95% of repeated samples, were this procedure to be repeated. The larger the trial, the smaller the 95% CI and the more certain the effect estimate.

Imprecision in the evidence reviews is assessed by considering whether the width of the 95% CI of the effect estimate is relevant to decision-making, taking each outcome in isolation. This assessment also involves effect size thresholds for clinical importance (the minimally important difference, MID) for benefit and for harm.

If the effect estimate CI includes clinically important benefit (or harm) there is uncertainty over which decision to make (based on this outcome alone). The CI is consistent with 2 possible decisions and so this is considered to be imprecise in the GRADE analysis and the evidence is downgraded by 1 level (‘serious imprecision’).

An effect CI including clinically important benefit, clinically important harm and no effect is consistent with 3 possible decisions. This is considered to be very imprecise in the GRADE analysis and the evidence is downgraded by 2 levels (‘very serious imprecision’).

Minimally important differences

The literature was searched for established MID for the selected outcomes in the evidence reviews. In addition, the committee was asked whether they were aware of any acceptable MID in the clinical community. See Table 6 for a list of the published MID used in this guideline.

If no published or acceptable MID were identified, the committee considered whether it was clinically acceptable to use the GRADE default MID to assess imprecision. For dichotomous outcomes clinically important thresholds for a RR or 0.8 and 1.25

respectively were used. For continuous outcomes, GRADE default MID are half of the SD of the control group at baseline, or if not available at follow up.

Table 6: MID reported in the literature for selected measures

Name of measure	Full name of measure	Conditions covered	Purpose	MID
ePAQ-PF	Electronic Patient Assessment Questionnaire – Pelvic Floor	Urinary/Bowel/Vaginal/Sexual	Symptoms + QoL	At 3 months follow up[a]: ±14.1 for OAB domain ±43.6 for SUI domain ±54.7 for urinary quality of life domain ±3.4 for prolapse domain
ICIQ-UI	International Consultation on Incontinence Modular Questionnaire – Urinary incontinence	OAB/SUI	Symptoms	±2.52 at 4-months follow up[b] ±5 at 1-year follow up[c] ±4 at 2-years follow up[c]
i-QOL	Urinary Incontinence Quality of Life Scale	OAB/SUI	QoL	At 12-weeks follow up[d]: ±2.5 between-treatment difference ±6.3 within-treatment difference
KHQ	Kings Health Questionnaire	OAB/SUI	Symptoms + QoL	±5 for OAB at 3-6 months follow up[e]
PFDI-20	Pelvic Floor Distress Inventory – Short Form	OAB/SUI/POP	Symptoms + QoL	±45 at 3-6 months follow up[f]
PFIQ-7	Pelvic Floor Impact Questionnaire – Short Form	OAB/SUI/POP	QoL	±36 at 3-6 months follow up[f]
PISQ	Pelvic Organ Prolapse/Incontinence Sexual Questionnaire	OAB/SUI/POP	Symptoms + QoL	±6 at 3 months follow up[g]
POP-SS	Pelvic Organ Prolapse Symptom Score	POP	Symptoms + QoL	±1.5 at 2 years follow up[h]
UDI	Urinary Distress Inventory	OAB/SUI	Symptoms	At 3 months follow up[i]: ±11.1 total score

Name of measure	Full name of measure	Conditions covered	Purpose	MID
				±7.5 for stress subscale

MID: Minimally Important Difference; OAB: Overactive Bladder; QoL: Quality of Life; POP: Pelvic Organ Prolapse; SUI: Stress Urinary Incontinence.
Notes: [a], Jones 2009; [b], Nyström 2015; [c], Sirls 2015; [d], Yalcin 2005; [e], Kelleher 2004; [f], Barber 2005; [g], Mamik 2014; [h], Hagen 2010; [i], Barber 2010.

Diagnostic test accuracy reviews

Modified GRADE methodology for diagnostic test accuracy reviews

The GRADE approach was modified to assess the quality of evidence about diagnostic test accuracy by adapting the principles of GRADE for intervention reviews as described below. Four domains were considered: risk of bias, indirectness, inconsistency and imprecision. Each domain was rated as 'no serious..', 'serious ..' or 'very serious ..' concerns. These domains were then combined to give the overall certainty in the body of evidence, rated as 'very low', 'low', 'moderate' or 'high'.

Assessing risk of bias in diagnostic test accuracy reviews

Risk of bias in diagnostic test accuracy studies was assessed using the risk of bias items from the QUADAS-2 checklist (see appendix H in Developing NICE guidelines: the manual 2014. An overall risk of bias judgement was for each study was reached by considering the QUADAS-2 bias domains together. The risk of bias for the body of diagnostic test accuracy evidence was based on the risk of bias from the individual studies but with consideration of how much each study contributed to the overall evidence base.

Assessing indirectness in diagnostic test accuracy reviews

Indirectness was assessed using the applicability items from the QUADAS-2 checklist. An overall indirectness judgement was for each study was reached by considering the QUADAS-2 applicability domains together. The indirectness for the body of diagnostic test accuracy evidence was based on the indirectness of the individual studies but with consideration of how much each study contributed to the overall evidence base.

Assessing inconsistency in diagnostic test accuracy reviews

Where there were multiple studies the body of evidence was downgraded for serious inconsistency if there was unexplained variability between studies, when viewed on a forest plot or Receiver Operating Characteristics (ROC) curve. If there was only one study then inconsistency was rated as 'not applicable'.

Assessing imprecision in diagnostic test accuracy reviews

Imprecision was judged by comparing the CI of the estimate of sensitivity or specificity to clinical decision thresholds agreed beforehand by the committee. The committee decided whether sensitivity or specificity was the most important for

decision making and agreed two threshold values. First a threshold for high sensitivity/specificity (above which the test would be definitely recommended) and second a threshold for low sensitivity/specificity (below which the test would not be recommended). If the CI of the estimate of sensitivity or specificity included one of these thresholds then the evidence was downgraded for serious imprecision, because it was consistent with two possible decisions. If the CI included both these thresholds then the evidence was downgraded for very serious imprecision because it was consistent with three possible decisions.

Qualitative reviews

GRADE CERQual methodology for qualitative reviews

The GRADE-CERQual (Confidence in the Evidence from Reviews of Qualitative research; Lewin 2015) approach was used to summarise the confidence in qualitative evidence. Each qualitative study was summarised by theme and meta-synthesis was carried out where appropriate to identify an overarching framework of themes and subthemes.

The overall confidence in evidence about each theme or sub-theme was rated as high, moderate, low or very low based on four dimensions: methodological limitations, applicability, coherence and adequacy of data.

Methodological limitations refer to the extent to which there were problems in the design or conduct of the studies that contributed evidence to the findings of the review.

Applicability of evidence was assessed by looking at the extent to which the body of evidence from the primary studies supporting the review findings is applicable to the review protocol

Coherence of findings was assessed by looking at the extent to which the review findings were well grounded in data from the contributing primary studies

Adequacy of data was assessed by looking at the degree of richness and quantity of data supporting the findings of the review

Assessing risk of bias in qualitative reviews

For qualitative studies, quality was assessed using a checklist for qualitative studies (as suggested in appendix H in Developing NICE guidelines: the manual 2014. This was based on the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies.

Evidence statements

Evidence statements are summary statements that are presented after the GRADE profiles highlighting the key features of the clinical evidence presented. The wording of the evidence statements reflects the certainty or uncertainty in the estimate of effect. The evidence statements are presented by outcome or theme and encompass the following key features of the evidence:

- the quality of the evidence (including GRADE rating, where relevant)

- the number of studies and/or the number of participants for a particular outcome (or theme in the case of qualitative evidence)
- a brief description of the participants
- the clinical significance of the effect and an indication of its direction (for example, if a treatment is clinically important (beneficial or harmful) compared with another, or whether there is no clinically important difference between the tested treatments).

Reviewing economic evidence

Systematic reviews of economic literature were conducted for all review questions covered in the guideline, unless economic evidence was not relevant to a review question. In addition, literature on the health-related quality of life of people covered by this guideline was systematically searched to identify studies reporting appropriate health state utility data that could be utilised in a cost-utility analysis.

Inclusion and exclusion of economic studies

The titles and abstracts of papers identified through the searches were independently assessed for inclusion using predefined eligibility criteria defined in Table 7.

Table 7: Inclusion criteria for the systematic reviews of economic evaluations

Inclusion criteria
Only studies from Organisation for Economic Co-operation and Development member countries were included, as the aim of the review was to identify economic information transferable to the UK context.
Selection criteria based on types of clinical conditions and population as well as interventions assessed were identical to the clinical review.
Only studies published from 2007 onwards were included in the review. This date restriction was imposed so that retrieved economic evidence was relevant to current healthcare settings and costs.
Studies were included provided that sufficient details regarding methods and results were available to enable the methodological quality of the study to be assessed, and provided that the study's data and results were extractable. Conference abstracts, poster presentations or dissertation abstracts were excluded.
Full economic evaluations (cost utility, cost effectiveness, cost benefit or cost consequence analyses) that assess both the costs and outcomes associated with the interventions of interest. Cost studies were also considered for the inclusion.

Once the screening of titles and abstracts was complete, full versions of the selected papers were acquired for assessment. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for the search of economic evaluations is presented in appendix D of this chapter.

Lists of included economic studies with their evidence tables, as well as studies excluded after obtaining full text with reasons for exclusion, are provided in appendix H and appendix K of the respective Evidence Review Reports.

Appraising the applicability and quality of economic evidence

The applicability and quality of economic evaluations in this guideline were appraised using the methodology checklist reported in the Developing NICE guidelines: the manual 2014, appendix M for all studies that met the inclusion criteria.

The methodological assessment of economic studies considered in this guideline has been summarised in economic evidence profiles that were developed for each review question for which economic evidence was available. All studies that fully or partially met the applicability and quality criteria described in the methodology checklist were considered during the guideline development process.

Health economic profiles of all economic studies that were considered during guideline development, including de novo economic analyses undertaken for this guideline, are provided in appendix I of the respective Evidence Review Reports.

Health economic modelling

The aims of the health economic input to the guideline were to inform the guideline committee of potential economic issues related to the management of women with stress urinary incontinence or pelvic organ prolapse in order to ensure that recommendations represented a cost-effective use of healthcare resources. Health economic evaluations aim to integrate data on healthcare benefits (ideally in terms of quality-adjusted life-years, QALYs) with the costs of different care options. In addition, the health economic input aimed to identify areas of high resource impact; recommendations which might have a large impact on Clinical Commissioning Group or Trust finances need to be supported by robust evidence on cost effectiveness.

Areas for economic modelling were prioritised by the committee. The rationale for prioritising review questions for economic modelling was set out in an economic plan agreed between NICE, the committee, and members of the Developer's technical team. Economic modelling was undertaken in areas with likely major resource implications, where the current extent of uncertainty over cost effectiveness was significant and economic analysis was expected to reduce this uncertainty. The following economic questions were selected as key issues that were addressed by economic modelling:

- cost effectiveness of surgical management options (including mesh and non-mesh procedures) for pelvic organ prolapse
- cost effectiveness of combined stress urinary incontinence and pelvic organ prolapse surgery to prevent postoperative urinary incontinence in women having surgery for pelvic organ prolapse

Also, the cost effectiveness of anticholinergic drugs for overactive bladder (with the focus on the risks to cognitive function) was prioritised for de-novo economic modelling. However, clinical data was insufficient to inform economic modelling in this area.

The methods and results of the de novo economic analyses are reported in appendix J of Evidence Reports of the respective review questions. When new economic analysis was not prioritised, the committee made a qualitative judgement regarding cost effectiveness by considering expected differences in resource use and costs

between options, alongside clinical effectiveness evidence identified from the clinical evidence review.

Cost effectiveness criteria

NICE's report Social value judgements: principles for the development of NICE guidance sets out the principles that committees should consider when judging whether an intervention offers good value for money. In general, an intervention was considered to be cost effective if any of the following criteria applied (given that the estimate was considered plausible):

- the intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or
- the intervention cost less than £20,000 per QALY gained compared with the next best strategy.

The committee's considerations of cost-effectiveness are discussed explicitly under the 'Cost effectiveness and resource use' headings of the relevant sections.

Developing recommendations

Guideline recommendations

Recommendations were drafted on the basis of the committee's interpretation of the available evidence, taking into account the balance of benefits, harms and costs between different courses of action. When clinical and economic evidence was of poor quality, conflicting or absent, the committee drafted recommendations based on the members' expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs or implications compared with the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues.

The main considerations specific to each recommendation are outlined under the 'Recommendations and link to evidence' headings within each Evidence Report.

For further details please refer to the Developing NICE guidelines: the manual 2014.

Research recommendations

When areas were identified for which good evidence was lacking, the committee considered making recommendations for future research. For further details please refer to the Developing NICE guidelines: the manual 2014.

Validation process

This guidance is subject to a 6-week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website at publication. For further details please refer to the Developing NICE guidelines: the manual 2014.

Updating the guideline

Following publication, and in accordance with the NICE guidelines manual, NICE will undertake a review of whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update. For further details please refer to the Developing NICE guidelines: the manual 2014.

Funding

The NGA was commissioned by NICE to develop this guideline.

References

Barber 2009

Barber MD, Spino C, Janz NK, Brubaker L, Nygaard I, Nager CW, Wheeler TL. Pelvic Floor Disorders Network. The minimum important differences for the urinary scales of the Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire. American Journal of Obstetrics and Gynecology, 200, 580-e1, 2009

Barber 2005

Barber MD, Walters MD, Bump RC. Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). American Journal of Obstetrics and Gynecology, 193, 103-13, 2005

Brazzelli 2018

Brazzelli M, Javanbakht M, Imamura M, Hudson J, Moloney E, Becker F, et al. The Effectiveness and cost-effectiveness of Surgical Treatments for womEn with stRess urinary incontinence: An evidence synthesis, economic evaluation and discrete choice experiment (ESTER). Health Technology Assessment, 2018, in review

DerSimonian 2015

DerSimonian R, Laird N. Met-analysis in clinical trials revisited. Contemporary Clinical Trials 45, 139-145, 2015.

Hagen 2010

Hagen S, Glazener C, Cook J, Herbison P, Toozs-Hobson P. Further properties of the pelvic organ prolapse symptom score: minimally important change and test-retest reliability. Neurourology and Urodynamics, 29, 1055-6, 2010

Lewin 2015

Lewin S, Glenton C, Munthe-Kaas H et al.Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). PLoS Med 12(10), e1001895 2015**Jones 2009**

Jones GL, Radley SC, Lumb J, Farkas A. Responsiveness of the electronic personal assessment questionnaire-pelvic floor (ePAQ-PF). International Urogynecology Journal, 20, 557-64, 2009

Kelleher 2004

Kelleher CJ, Pleil AM, Reese PR, Burgess SM, Brodish PH. How much is enough and who says so? The case of the King's Health Questionnaire and overactive bladder. BJOG: An International Journal of Obstetrics & Gynaecology, 111, 605-12, 2004

Mamik 2014

Mamik MM, Rogers RG, Qualls CR, Morrow JD. The minimum important difference for the pelvic organ prolapse-urinary incontinence sexual function questionnaire. International Urogynecology Journal, 25, 1321-6, 2014

NICE 2013

National Institute for Health and Care Excellence (NICE) (2013) Urinary incontinence in women: management. Clinical guideline [CG171] (updated 2015). Available from https://www.nice.org.uk/guidance/CG171 (accessed 28th August 2018).

NICE 2014

National Institute for Health and Care Excellence (NICE) (2014) Developing NICE guidelines: the manual (updated 2017). Available from https://www.nice.org.uk/process/pmg20/chapter/introduction-and-overview (accessed 1st August 2018)

NICE 2018

National Institute for Health and Care Excellence (NICE) (2014) NICE Policy on conflict of interest (updated 2017). Available from https://www.nice.org.uk/Media/Default/About/Who-we-are/Policies-and-procedures/declaration-of-interests-policy.pdf (accessed 1st August 2018)

Nyström 2015

Nyström E, Sjöström M, Stenlund H, Samuelsson E. ICIQ symptom and quality of life instruments measure clinically relevant improvements in women with stress urinary incontinence. Neurourology and Urodynamics, 34, 747-51, 2015

Santesso 2016

Santesso N, Carrasco-Labra A, Langendam M, Brignardeloo-Petersen R, Mustafa RA, Heus P, Lasserson T, Opiyo N, Kunnamo I, Sinclair D, Garner P, Treweek S, Tovey D, Akl EA, Tugwell P, Brozek JL, Guyatt G, Schunemann HJ. Improving GRADE evidence tables part 3: detailed guidance for explanatory footnotes supports creating and understanding GRADE certainty in the evidence judgments. Journal of clinical epidemiology 74, 28-39 2016

Sirls 2015

Sirls LT, Tennstedt S, Brubaker L, Kim HY, Nygaard I, Rahn DD, Shepherd J, Richter HE. The minimum important difference for the International Consultation on Incontinence Questionnaire—Urinary Incontinence Short Form in women with stress urinary incontinence. Neurourology and Urodynamics, 34, 183-7, 2015

Yalcin 2006

Yalcin I, Patrick DL, Summers K, Kinchen K, Bump RC. Minimal clinically important differences in Incontinence Quality-of-Life scores in stress urinary incontinence. Urology, 67, 1304-8, 2006