PEOPLE study

Pessary or Surgery for a Symptomatic Pelvic Organ Prolapse

Supplement 2: Statistical Analysis Plan

Table of contents

1. Statistical analysis plans

Final statistical analysis plan: Version 2, 9th April 2022

1.2 Final statistical analysis plan

PEOPLE study

Pessary or Surgery for a Symptomatic Pelvic Organ Prolapse

Statistical Analysis Plan

Version 2

EUDRA CT no. **Netherlands Trial Register**	n/a NL4756
Principal investigator	Dr. A. Vollebregt
Coordinating investigator	Prof. Dr. C.H. van der Vaart
Sponsor	University Medical Center Utrecht
Initial date SAP	09-07-2021
Revision Date	09-04-2022
Trial methodologist	Dr. R.G. Duijnhoven
Author	Drs. L.R. van der Vaart

Table of contents

1. SAP signatures

I give my approval for the attached SAP entitled *Pessary or Surgery for a Symptomatic Pelvic Organ Prolapse* dated Version 2, 09 April 2022.

Primary investigator

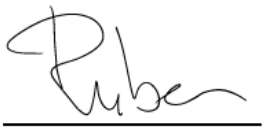
Name: Dr. A. Vollebregt

Signature: 

Date: 10-05-2022

Trial methodologist

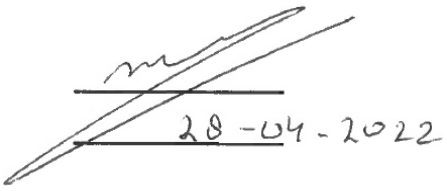
Name: Dr. R.G. Duijnhoven

Signature: 

Date: 2 May 2022

Coordinating investigator

Name: Prof. Dr. C.H. van der Vaart

Signature: 

Date: 28-04-2022

2. List of abbreviations

CRF	Case Record Form
DSMB	Data Safety Monitoring Board
EMA	European Medicines Agency
EudraCT	European Clinical Trials Database
FSF	Female sexual functioning
GP	General Practitioner
IQR	Interquartile range
ITT	Intention-to-treat
METC	Research ethics committee / *Medisch Ethische Toetsingscommissie*
NSA	Not sexually active
PFDI-20	Pelvic Floor Distress Inventory
PFIQ-7	Pelvic Floor Impact Questionnaire
PGI-I	Patient Global Impression of Improvement scale
PGI-S	Patient Global Impression of Severity scale
PISQ-IR	Pelvic Organ Prolapse / Urinary Incontinence Sexual Questionnaire
POP	Pelvic Organ Prolapse
PP	Per-protocol
QoL	Quality of life
SA	Sexually active
SAP	Statistical Analysis Plan
SD	Standard Deviation
WMO	*Wet Medisch-wetenschappelijk Onderzoek met Mensen*

3. Changes compared to previous version

The current version is the second version.

Version	Date	Revision Details	Prepared	Checked	Approved
D	05 Oct 2021	Draft Issue	LvV	RD	n/a
1	02 Apr 2022	First Issue	RD / LvV	AV / CHV	AV / CHV
2	09 Apr 2022	Clarification multiple imputation	RD / LvV	AV / CHV	AV / CHV

4. Introduction

The full background and rationale for the trial can be found in the PEOPLE trial protocol. In summary, female pelvic organ prolapse (POP) is a common problem in women that negatively affects quality of life due to micturition and defecatory symptoms, sexual disorders and vaginal bulging [1]. The estimated prevalence of symptomatic POP among women between 45-85 years of age is 8.3 – 23.7% and the lifetime risk of undergoing surgery for POP is 20% by the age of 80 [2-6].

It is current practice in the Netherlands that the general practitioner (GP) treats the majority of women with POP symptoms. Women with moderate to severe POP symptoms are often referred to a gynecologist for treatment. This study focuses at the subgroup of moderate to severe POP.

Known effective treatment options for moderate to severe POP are pessary and surgery. Studies regarding pessary for this indication however, are mainly observational in nature and inherently subject to selection and indication bias [7]. In the literature, outcomes of pessary therapy are mainly recorded in terms of (dis-) continuation of therapy and to a much lesser extent in terms of symptom relief. Although pessary therapy is minimally invasive, side effects may occur in up to 56% of women and include vaginal bleeding, pessary expulsion, excessive vaginal discharge, pain and urinary- or fecal incontinence [8]. Side effects, among the wish for surgical intervention, can be the reason for discontinuation of pessary therapy. In current literature, the discontinuation rate varies between 21.8 – 36% at 24-months follow-up [9, 10]. Pessary therapy however, is inexpensive and costs are mainly related to doctor visits and treatment of side effects. In case of self-management costs might even be lower.

Surgery for POP results in much to very much relieve of symptoms in 80% of women and achieves significant improvement of quality of life [11-13]. Side effects of POP surgery can include temporary urinary retention, temporary buttock pain in case of sacrospinous hysteropexy, hematoma, urinary tract infection, newly reported dyspareunia (10%), *de novo* urinary stress incontinence (9.9%), recurrence of POP (36% in 10 years follow-up) and a reoperation rate of 17% [14-18]. These complications seldom lead to persistent morbidity.

Although clinical efficacy appears to favor surgery, the large variation in study design, outcome measures and loss to follow up makes any conclusions on the best treatment option speculative. Therefore, there is an urgent need for a long-term comparative study between pessary and surgery, which has been recognized by reviews [7, 19, 20]. Based on current cohort and case-control studies we hypothesize that a strategy of initial pessary therapy for moderate to severe POP, is more cost-effective than surgery.

The objective of the PEOPLE trial is to investigate the non-inferiority of effectiveness, and cost-effectiveness of pessary therapy for moderate to severe POP, as compared to surgery.

5. Patients

5.1 Study population

Women with a symptomatic POP, referred to the outpatient clinic by their GP.

5.2 Inclusion criteria

- Women with a prolapse stage 2 or more
- Women with moderate to severe POP symptoms. Moderate to severe POP symptoms is defined as a prolapse domain score > 33 on the validated Dutch version of the Urogenital Distress Inventory [21]
- Women who have had a successful pessary fitting procedure
- Written informed consent

5.3 Exclusion criteria

- Prior urogynaecological (prolapse or incontinence) surgery
- Probability of future childbearing
- Insufficient knowledge of the Dutch language
- Co-morbidity causing increased surgical risks at the discretion of the surgeon
- Major psychiatric illness
- Prior pessary use

6. Study design

Multicenter randomized controlled non-inferiority trial comparing pessary therapy as the intervention to surgery as the standard, including an economic evaluation. The economic evaluation is outside the scope of this Statistical Analysis Plan (SAP).

7. Randomization and blinding

7.1 Randomization procedure

Prior to randomizing patients, all eligible women who give consent for participating in the trial, will have a short (30 min.) pessary fitting. This ensures that only women who are fit for both treatment options enter the randomization procedure. The trial fitting intends to prevent any patients for whom pessary cannot be fitted being randomized. This trial fitting is not intended for, or to asses symptom relief.

Baseline characteristics of those women with an unsuccessful pessary fitting will be recorded to allow analyses of this group.

When fitting is successful, women will be randomly allocated in a 1:1 ratio to either treatment with a pessary or surgical treatment. Randomization will be done using the web-based software package ALEA. The randomization sequence will be computer generated using variable blocks of sizes two and four. Randomization is stratified by center.

7.2 Blinding

Due to the nature of the intervention masking of the intervention is not possible.

8. Ethical approval and consent

8.1 Ethical approval

This study will be conducted according to the principles of the Declaration of Helsinki (version 10, October 2013) and in accordance with the Medical Research Involving Human Subjects Act (WMO) and other guidelines, regulations and Acts.

This study was approved by the Medical research ethics committee (MREC); in Dutch: medisch ethische toetsing commissie (METC). The date of approval was 22nd September 2015 and the METC protocol number is 14-533/M. The METC required a Data Safety Monitoring Board (DSMB). The monitoring was coordinated by the Dutch Consortium and was executed by a qualified intern monitor. This person was not involved in design and output of this research. The frequency of checking was every year. The investigator will submit a summary of the progress of the trial to the accredited METC one a year.

8.2 Consent

Written informed consent is obtained from any patient before enrolment.
Any patients counselled for participation in the trial who refused to participate will be anonymously registered in the online randomization program ALEA.

9. Treatment of subjects

9.1 Pessary therapy

All pessaries used for treatment in this trial need to be made of modern silicon material. Any such types of pessaries, both supportive and occlusive/space filling are allowed. A choice of pessary is to be made according to the treating gynecologists judgement. After placing the pessary, all women will receive verbal and written instructions on the self-management of pessary therapy. Patients may either continue supervised pessary management by their GP or gynecologist, or self-manage pessary care.

The first follow-up visit after pessary placement will always be performed by the gynecologist. In case of self-managed pessary treatment, the frequency of cleaning is left to her personal judgment, but may not exceed 4 months. If self-management is not possible or not preferred, women will be seen at 4 months intervals for pessary cleaning and vaginal

inspection, preferable by their GP. In case of vaginal atrophy, the use of topical estrogens will be advised in accordance with the guidelines. The diagnosis of atrophy is left to the judgment of the treating physician.

9.2 Surgical intervention

All surgical procedures will be performed according to national guidelines. The decision which technique to use is left, to the discretion of the gynecologist, within the limitations below [22].

Cystocele repair will consist of conventional anterior colporrhaphy [23]. For uterine descent different techniques are allowed [14]. These techniques can either be uterus sparing (sacrospinous hysteropexy [24], modified Manchester-Fothergill procedure [25] or an abdominal sacrocolpopexy [23]) or a vaginal hysterectomy.

A coexistent stage 2 rectocele repair will be a conventional colporrhaphy posterior. All procedures are performed under general or spinal anesthesia and under antibiotics and thrombosis prophylaxis according to local protocols.

10. Objectives

10.1 Primary objective

To assess if pessary treatment is therapeutically non-inferior to surgical intervention at 24-months follow-up.

10.2 Secondary objectives

- Type and number of complications and re-interventions for both treatments after 24-months
- Changes in symptom bother and disease-specific quality of life at 24-months
- Changes in subjective severeness of symptoms at 24-months
- Changes in sexual function at 24-months follow-up

11. Endpoints and outcome measures

11.1 Primary outcome measures

The primary endpoint of therapeutic efficacy is measured using the Patient Global Impression of Improvement of POP symptoms at 24-months by means of the PGI-I on a 7-point scale [26].

Global impression of improvement scores will be dichotomized as either successful or unsuccessful improvement. Success is defined as reported ‘much’ or ‘very much’ improvement. No success is defined as a ‘little better’, ‘no change’, ‘a little worse’, ‘much worse’ or ‘very much worse’, on the PGI-I.

Inferences on non-inferiority will be made using the risk difference against the non-inferiority margin (see also Ch. 12 *Sample size and non-inferiority margin*).

11.2 Secondary outcome measures

Adverse outcomes will be measured as proportions of dichotomous outcomes. Outcome definitions are according to clinical practice.

Re-interventions include switching from pessary to surgery, re-surgery after initial surgery or additional use of a pessary after surgery.

Symptom bother and disease-specific quality of life are (QoL) measured with the Pelvic Floor Distress Inventory (PFDI-20) and Pelvic Floor Impact Questionnaire (PFIQ-7) respectively [27, 28]. The PFDI-20 and PFIQ-7 are continuous outcomes and the delta of change between baseline and follow-up will be measured. The subscale scores for both questionnaires vary from 0-100 and the total scores from 0-300 [28]. Negative values in the delta of change represent improvement, whereas positive values indicate deterioration.

Subjective severeness of symptoms is measured with the Patient Global Impression of Severity scale (PGI-S), the outcomes are ordinal and measured using interval scales. Reduction of at least one point from baseline was considered as success.

Sexual function is measured with the Pelvic Organ Prolapse / Urinary Incontinence Sexual Questionnaire, IUGA-Revised (PISQ-IR) distinguishing sexually active (SA) and inactive women (NSA), this is a continuous outcome and the delta of change between baseline and follow-up will be measured [29]. The PISQ-IR comprises of 6 domain-specific subscales for SA women, where a higher score indicates better female sexual functioning (FSF), and 4 for NSA women, where a higher score indicates a greater impact of POP on sexual inactivity [29]. For SA women, a summary score was calculated to provide an overall effect on FSF in assessing the clinical management of POP [30]. Each domain has a minimum score of 1 and maximum score of either 4 or 5. For SA women, an increase in the delta of change indicates less impact on FSF and better sexual functioning. For NSA women, a decrease in the delta of change indicates less impact of POP on sexual inactivity [29].

12. Sample size and non-inferiority margin

With 198 women per group (396 patients in total), there is 80% power to reject the null hypothesis that pessary therapy is inferior to surgery, with a one-sided alpha of 0.05, a non-inferiority margin of 10% and an incidence of successful improvement measured using the Global impression of improvement of 80% in the group treated surgically. The sample size was determined using the Z-test with unpooled variances.

Allowing for an attrition rate of up to 10% loss of follow-up, a total of 436 patients will be recruited.

13. Data sources and Data Quality

Three data sources will be used for the study, all of which are collected digitally.

13.1 Randomization data

Compliance of the eligible patient with in- and exclusion will be recorded in the randomization database, prior to randomization. The ALEA randomization software records the in- and exclusion details, patient age at randomization, and assigns and records the patient study record identification number. Treatment allocation is recorded unblinded and is visible to all staff with access to the system.
The ALEA randomization software records changes made to the data entered in an audit trail.

13.2 Case Report Form database

A database containing the data from case record forms (CRF) is developed in OpenClinica (electronic CRF, eCRF simply referred to as CRF). All data will be collected digitally. On site monitoring will be conducted to assess overall study compliance and to conduct source data verification. Programmed cross checks and checks for value ranges are applied to facilitate correct data entry. Central monitoring strategies will be used to track study progress, and progress and quality of data entry.

13.3 Questionnaires

Questionnaires will be sent to study participants electronically using Limesurvey 2.6.7 (forums.limesurvey.org). Questionnaire responses will also be stored in Limesurvey.

14. General outline of analysis

14.1 Trial profile

The flow of study participants will be displayed using the CONSORT-statement model diagram (see Ch. 18.1 *CONSORT flow diagram of study population*).

14.2 Analysis populations

Full Analysis Population
The full analysis population (intention-to-treat, ITT) is defined as the population of all patients randomized.

Per-protocol population
The per-protocol population (PP), is defined as the group of women who started surgery of pessary treatment as randomized, but excludes any women who did not initiate treatment as randomized, those who discontinued use of pessary or underwent (re-)surgery for treatment of POP complaints within 24-months from baseline, and those who started using a pessary after initially being allocated to and having had surgery.

Safety population
The safety population, reporting patients by treatment received will not be used.

15. Planned analysis

15.1 Baseline characteristics
Baseline characteristics will be presented as descriptive statistics for both treatment groups as numbers with percentages, or as averages (mean or median) with standard deviations (SD) or interquartile ranges (IQR) as appropriate.

A mock table is given in **17.1.**

15.2 Primary outcome
Primary efficacy analysis
The primary outcome is the dichotomized subjective improvement, measured with the PGI-I at 24-months follow-up on the intention-to-treat population. Non-inferiority will be assessed using the risk difference between surgery (ref.) and pessary. The 90% confidence interval should not exceed the non-inferiority limit of 10% for non-inferiority to be proven. A generalized linear model with binomial distribution and identity link will be used to determine the point estimate and confidence interval while adjusting for the stratification by site in the randomization. Inferential testing to obtain a p-value will be done using the Farrington-Manning test (unadjusted).

The null hypothesis assumes that pessary therapy is inferior to surgery, the alternative hypothesis assumes that pessary therapy is non-inferior to surgery.

A mock table is given in **17.4.**

Secondary efficacy analyses
As a secondary efficacy analysis, the primary outcome will be estimated on the per-protocol population. Inferential testing will be done as for the primary efficacy analysis.

A mock table is given in **17.4.**

15.3 Secondary outcomes
Complications, re-interventions and procedural details will be presented as absolute numbers with percentages. Relative measures of effect will be used to compare groups (relative risks), with Chi-squared tests or Fisher's exact test as appropriate.

A mock table is given in **17.5**.

The cross-over from pessary to surgical intervention will be presented as absolute number with percentages. Also, a Kaplan-Meier plot will be presented for the time to re-intervention (switch to surgery for the pessary group, re-surgery or additional use of a pessary for the

surgery group). The differences between survival curves will be assessed using a log-rank test.

Other secondary outcomes are the symptom bother (PFDI-20), disease-specific quality of life (PFIQ-7), subjective severeness of symptoms (PGI-S), and sexual function (PISQ-IR). Secondary outcomes were compared within groups and between groups at 24-months follow-up. The outcomes are evaluated as differences in means with 95% confidence intervals. If distributions are approximately normal, a t-test will be used, either with pooled or unpooled (Satterthwaite) variances. If measurements are not considered normally distributed, either confidence intervals will be estimated by means of bootstrapping, or non-parametric methods will be used.

Secondary outcomes will be assessed on both the ITT as well as the per-protocol populations.

A mock table is given in **17.4.**

15.4 Subgroup analysis

Subgroup analyses for the main outcome will be conducted for exploratory purpose on sexual status (sexually active or sexually inactive). Subgroup analysis will be conducted in a generalized linear model with treatment group and sexual status together with an interaction term in the regression model.

15.5 Baseline differences

Primary analyses of the main outcome will not be adjusted for baseline characteristics. In case of potentially important baseline differences, exploratory adjusted analyses can be considered *post-hoc*.

15.6 Missing data

Notwithstanding the choice of imputating missing data, missing data patterns and drop-out will be assessed and reported.

Missing primary outcome data will be imputated if results before imputation indicate non-inferiority. If non-inferiority is not shown imputation will be conducted provided the missingness and dropouts do not favor the alternative hypothesis, as stipulated in European Medicines Agency (EMA) guidance [31].

If non-inferiority is shown, multiple imputation will be used to assess the effect of missing data on the observed non-inferiority. Objective would be to decrease the risk of falsely rejecting the null hypothesis.

Missing data of secondary outcomes will not be imputated.

15.7 Interim analysis and data monitoring

A Data Safety Monitoring Board (DSMB) is established for this study to monitor safety. A formal interim analysis for efficacy will not be conducted.

15.8 Multiple testing

Given that there is only one primary outcome and testing is only to be done at 24-months follow-up, adjustment for multiple testing will not be made.

16. Presentation of study results

Details on recruitment and treatment compliance are included in the flow diagram of participants.
Time to switching of therapy, or initiation of additional therapy will be presented in a Kaplan-Meier plot.

Mock tables for the primary outcome and secondary outcomes are included in the following chapter of this SAP. These cover general baseline characteristics, condition specific baseline characteristics upon physical examination, details on the intervention initiated, main efficacy measures, change in sexual status, complications and side effects.

17. Mock tables

17.1 Baseline characteristics

Table 17.1. Baseline characteristics of the study population		
Baseline characteristic	Pessary group (*n*=NNN)	Surgery group (*n*=NNN)
Age (yr)	xx.x (x.x)	xx.x (x.x)
BMI (kg/m²)		
Mean (SD)	xx.x (x.x)	xx.x (x.x)
Obese (BMI >30)	NNN (%)	NNN (%)
Race		
Caucasian	NNN (%)	NNN (%)
Sub-Saharan African	NNN (%)	NNN (%)
Afro-Caribbean	NNN (%)	NNN (%)
Hindu-Caribbean	NNN (%)	NNN (%)
Middle-Eastern	NNN (%)	NNN (%)
Asian	NNN (%)	NNN (%)
Unknown	NNN (%)	NNN (%)
Smoking	NNN (%)	NNN (%)
Diabetes	NNN (%)	NNN (%)
Chronic pulmonary disease	NNN (%)	NNN (%)
Parity	x (x - x)	x (x - x)
Mode of delivery		
Caesarean section	NNN (%)	NNN (%)
Vacuum assisted delivery	NNN (%)	NNN (%)
Forceps delivery	NNN (%)	NNN (%)
3rd/4th degree perineal tear	NNN (%)	NNN (%)
Menopausal state		
Pre-menopausal	NNN (%)	NNN (%)
Postmenopausal	NNN (%)	NNN (%)
History of gynecological surgery	NNN (%)	NNN (%)
Uterus extirpation	NNN (%)	NNN (%)
Family history of prolapse	NNN (%)	NNN (%)
Anti-depressants	NNN (%)	NNN (%)
Duration of complaints (mths.)	x.x (x.x)	x.x (x.x)
Vaginal atrophy	NNN (%)	NNN (%)
Prolapse stage		
II	NNN (%)	NNN (%)
≥III	NNN (%)	NNN (%)
PGI-S		
I	NNN (%)	NNN (%)
II	NNN (%)	NNN (%)
III	NNN (%)	NNN (%)
IV	NNN (%)	NNN (%)
PFDI-20 domain score		
UDI-6	xx.x (x.x)	xx.x (x.x)
CRADI-8	xx.x (x.x)	xx.x (x.x)
POPDI-6	xx.x (x.x)	xx.x (x.x)
PFDI-20 total score	xx.x (x.x)	xx.x (x.x)
PFIQ-7 domain score		
UIQ-7	xx.x (x.x)	xx.x (x.x)
CRAIQ-7	xx.x (x.x)	xx.x (x.x)
POPIQ-7	xx.x (x.x)	xx.x (x.x)
PFIQ-7 total score	xx.x (x.x)	xx.x (x.x)
PISQ-IR sexually active		
Partner related	xx.x (x.x)	xx.x (x.x)

Condition specific	xx.x (x.x)	xx.x (x.x)
Global quality	xx.x (x.x)	xx.x (x.x)
Condition impact	xx.x (x.x)	xx.x (x.x)
Arousal – orgasm	xx.x (x.x)	xx.x (x.x)
Desire	xx.x (x.x)	xx.x (x.x)
Summary score	xx.x (x.x)	xx.x (x.x)
PISQ-IR sexually inactive		
Partner related	xx.x (x.x)	xx.x (x.x)
Condition specific	xx.x (x.x)	xx.x (x.x)
Global quality	xx.x (x.x)	xx.x (x.x)
Condition impact	xx.x (x.x)	xx.x (x.x)
Data are n (%), mean (sd), or median (IQR). Percentages based on the number of observations available.		

17.2 Physical examination

Table 17.2 Physical examination at baseline		
Physical examination	Pessary group (*n*=NNN)	Surgery group (*n*=NNN)
Examination:	NNN (%)	NNN (%)
Supine	NNN (%)	NNN (%)
Standing	NNN (%)	NNN (%)
Both	NNN (%)	NNN (%)
POP-Q (cm)		
Aa	x.x (x.x)	x.x (x.x)
Ba	x.x (x.x)	x.x (x.x)
C	x.x (x.x)	x.x (x.x)
HG	x.x (x.x)	x.x (x.x)
PB	x.x (x.x)	x.x (x.x)
TVL	x.x (x.x)	x.x (x.x)
Ap	x.x (x.x)	x.x (x.x)
Bp	x.x (x.x)	x.x (x.x)
D	x.x (x.x)	x.x (x.x)
Vulvar deviations	NNN (%)	NNN (%)
Malignancy	NNN (%)	NNN (%)
Lichen	NNN (%)	NNN (%)
Positive stress test		
Yes	NNN (%)	NNN (%)
No	NNN (%)	NNN (%)
Not performed	NNN (%)	NNN (%)
Data are n (%), mean (SD), or median (IQR). Percentages based on the number of observations available.		

17.3 Treatment details at initiation

Table 17.3 Treatment details at initiation		
	Pessary group (*n*=NNN)	Surgery group (*n*=NNN)
Initiated treatment as randomised	NNN (%)	NNN (%)
Time between randomization and treatment (days)	xx.x (x.x)	xx.x (x.x)
Pessary type:		
Supportive	NNN (%)	n/a
Occlusive	NNN (%)	n/a
Pessary self-managed	NNN (%)	n/a
If not:		
Unable to	NNN (%)	n/a
Preference	NNN (%)	n/a
Topical oestrogens:		
Starting	NNN (%)	NNN (%)
Continuation	NNN (%)	NNN (%)
Unknown	NNN (%)	NNN (%)
Surgery type:		
Anterior colporraphia	n/a	NNN (%)
Posterior colporraphia	n/a	NNN (%)
Sacrospinous hysteropexy	n/a	NNN (%)
Modified Manchester-Forthergill procedure	n/a	NNN (%)
Vaginal hysterectomy with McCall or SSF	n/a	NNN (%)
Laparoscopic sacrocolpopexy	n/a	NNN (%)
Laparoscopic sacrohysteropexy	n/a	NNN (%)
Laparoscopic sacrocervixopexy	n/a	NNN (%)
Retropubic sling	n/a	NNN (%)
Transobturator sling	n/a	NNN (%)
Data are n (%), mean (sd), or median (IQR). Percentages based on the number of observations available.		

17.4 Main outcomes at 24-months

Table 17.4. Intention-to-treat and per-protocol analysis of the primary and secondary outcomes at 24 months.

	Intention-to-treat analysis				Per-protocol analysis			
	Pessary group (*n*=NNN)	Surgery group (*n*=NNN)			Pessary group (*n*=NNN)	Surgery group (*n*=NNN)		
			Risk difference (90% CI)	*p*-value			Risk difference (90% CI)	*p*-value
PGI-I: improvement – no./total no. *	NNN/NNN (%)	NNN/NNN (%)	x.x (x.xx – x.xx)	0.xx	NNN/NNN (%)	NNN/NNN (%)	x.x (x.xx – x.xx)	0.xx
PGI-S: improvement – no./total no. ¥	NNN/NNN (%)	NNN/NNN (%)	x.x (x.xx – x.xx)	0.xx	NNN/NNN (%)	NNN/NNN (%)	x.x (x.xx – x.xx)	0.xx
Change in PFDI-20 domain score †			Mean difference (95% CI)				Mean difference (95% CI)	
POPDI-6	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
CRADI-8	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
UDI-6	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
PFDI total score	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Change in PFIQ-7 domain score †								
UIQ-7	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
CRAIQ-7	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
POPIQ-7	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
PFIQ-7 total score	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Change in PISQ-IR domain score SA ¤								
Partner related	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Condition specific	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Global quality	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Condition impact	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Arousal – orgasm	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Desire	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Summary score	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Change in PISQ-IR domain score NSA §								
Partner related	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Condition specific	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Global quality	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx
Condition impact	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx	xx.x (x.x)	xx.x (x.x)	x.x (x.xx – x.xx)	0.xx

***Farrington-Manning test for non-inferiority against the non-inferiority margin of -10%.**
¥ Chi-square test, risk difference 95%
† Mean difference with paired t-test. A negative change indicates improvement. Subscale scores vary from 0 – 100, total scores from 0 – 300.
¤ Mean difference with paired t-test. An increase in the delta of change indicates less impact on FSF and better sexual functioning.
§ Mean difference with paired t-test. A decrease in the delta of change indicates less impact of POP on sexual inactivity.

17.5 Adverse events and additional therapy

Table 17.5. Adverse events and additional therapy		
	Pessary group (*n*=NNN)	Surgery group (*n*=NNN)
Discontinuation pessary		n/a
12 months	NNN (%)	
24 months	NNN (%)	
Switch to surgery or add. therapy after surgery	NNN (%)	NNN (%)
Reason to switch to surgery of add. therapy after surgery		
Inadequate symptom relief	NNN (%)	NNN (%)
Recurrence of prolapse	NNN (%)	NNN (%)
Incontinence	NNN (%)	NNN (%)
Problems with sexual functioning	NNN (%)	NNN (%)
Discomfort / pain	NNN (%)	NNN (%)
Pessary expulsion	NNN (%)	NNN (%)
Excessive discharge	NNN (%)	NNN (%)
Dissatisfied with pessary self-management	NNN (%)	NNN (%)
Other	NNN (%)	NNN (%)
Adverse events		
Infection	NNN (%)	NNN (%)
Urinary tract infection	NNN (%)	NNN (%)
Urinary retention	NNN (%)	NNN (%)
Blood loss	NNN (%)	NNN (%)
Haematoma	NNN (%)	NNN (%)
Re-intervention	NNN (%)	NNN (%)
Other	NNN (%)	NNN (%)
Data are n (%). Percentages based on the number of observations available		

17.6 Change of sexual status

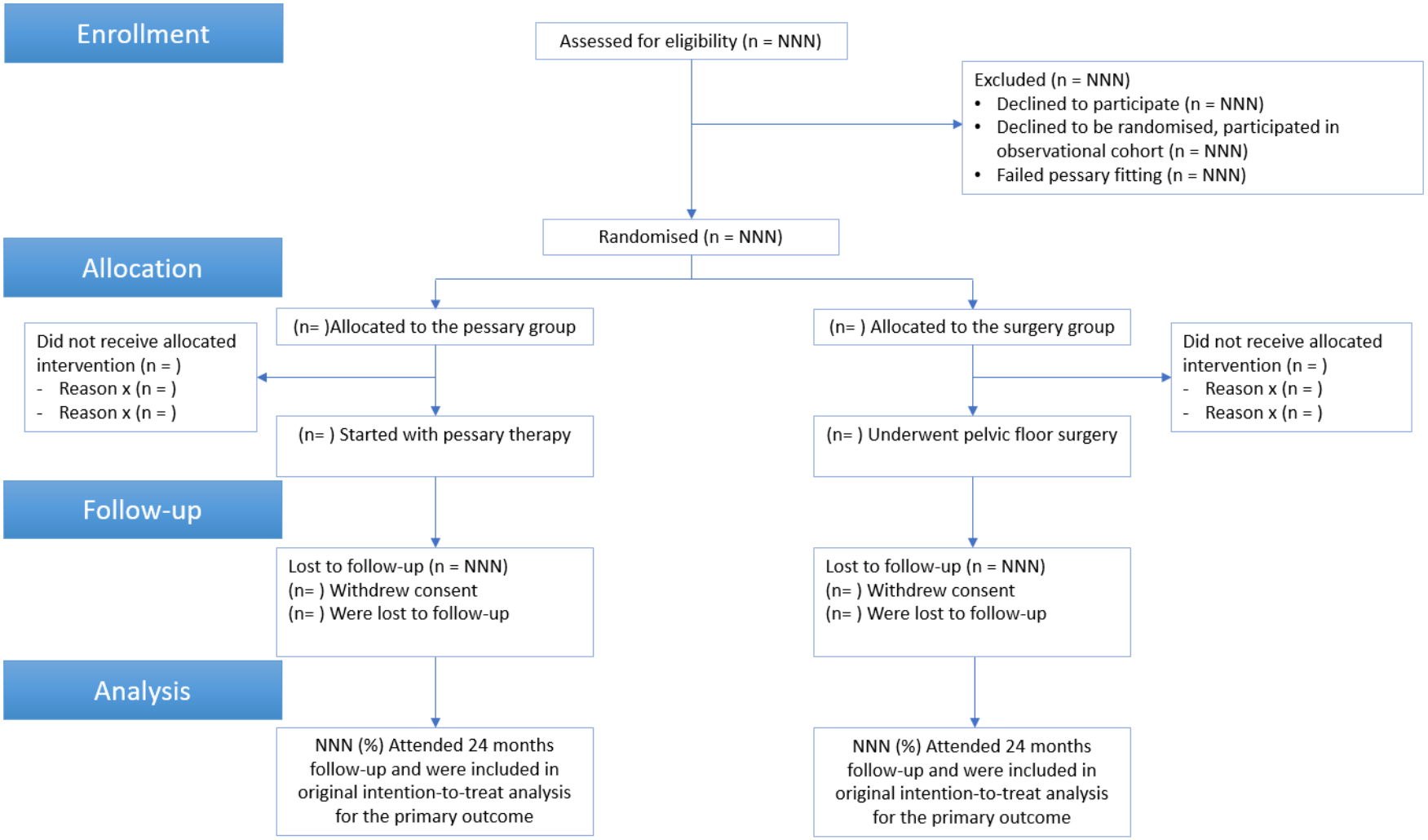
Table 17.6. Change of sexual status within 24 months.			
	Pessary group (*n*=NNN)	Surgery group (*n*=NNN)	Relative risk (95% CI), *p*-value
NSA at baseline			
Remained NSA	NNN (%)	NNN (%)	
Change from NSA to SA	NNN (%)	NNN (%)	x.xx (x.x – x.x), 0.xx
SA at baseline			
Remained SA	NNN (%)	NNN (%)	
Change from SA to NSA	NNN (%)	NNN (%)	x.xx (x.x – x.x), 0.xx
Data are n (%). Percentages based on the number of observations available.			

17.7 Subgroup analysis

Subgroup analyses will be reported using forest plots, and include the p-value for interaction.

18. Figures

18.1 CONSORT flow diagram of study population



18.2 Kaplan-Meier plots for time to intervention

Kaplan Meier plot of time to switch to first (surgical) re-intervention (treatment other than treatment allocated by randomization).

19. References

1. Jelovsek JE, Maher C and Barber MD. Pelvic organ prolapse. *The Lancet* 2007; 369: 1027-1038.
2. Mou T, Warner K, Brown O, et al. Prevalence of pelvic organ prolapse among US racial populations: A systematic review and meta-analysis of population-based screening studies. *Neurourol Urodyn* 2021; 40: 1098-1106.
3. Slieker-ten Hove MC, Pool-Goudzwaard AL, Eijkemans MJ, Steegers-Theunissen RP, Burger CW and Vierhout ME. The prevalence of pelvic organ prolapse symptoms and signs and their relation with bladder and bowel disorders in a general female population. *Int Urogynecol J Pelvic Floor Dysfunct* 2009; 20: 1037-1045.
4. Nygaard I, Barber MD, Burgio KL, et al. Prevalence of symptomatic pelvic floor disorders in US women. *JAMA* 2008; 300: 1311-1316.
5. Wu JM, Matthews CA, Conover MM, Pate V and Jonsson Funk M. Lifetime risk of stress urinary incontinence or pelvic organ prolapse surgery. *Obstet Gynecol* 2014; 123: 1201-1206.
6. Tegerstedt G, Maehle-Schmidt M, Nyrén O and Hammarström M. Prevalence of symptomatic pelvic organ prolapse in a Swedish population. *Int Urogynecol J Pelvic Floor Dysfunct* 2005; 16: 497-503.
7. Lamers BH, Broekman BM and Milani AL. Pessary treatment for pelvic organ prolapse and health-related quality of life: a review. *Int Urogynecol J* 2011; 22: 637-644.
8. Sarma S, Ying T and Moore KH. Long-term vaginal ring pessary use: discontinuation rates and adverse events. *BJOG* 2009; 116: 1715-1721.
9. Bugge C, Hagen S and Thakar R. Vaginal pessaries for pelvic organ prolapse and urinary incontinence: a multiprofessional survey of practice. *Int Urogynecol J* 2013; 24: 1017-1024.
10. Clemons JL, Aguilar VC, Sokol ER, Jackson ND and Myers DL. Patient characteristics that are associated with continued pessary use versus surgery after 1 year. *Am J Obstet Gynecol* 2004; 191: 159-164.
11. Lukacz ES, Warren LK, Richter HE, et al. Quality of Life and Sexual Function 2 Years After Vaginal Surgery for Prolapse. *Obstet Gynecol* 2016; 127: 1071-1079.
12. Vollebregt A, Fischer K, Gietelink D and van der Vaart CH. Primary surgical repair of anterior vaginal prolapse: a randomised trial comparing anatomical and functional outcome between anterior colporrhaphy and trocar-guided transobturator anterior mesh. *BJOG* 2011; 118: 1518-1527.
13. Withagen MI, Milani AL, den Boon J, Vervest HA and Vierhout ME. Trocar-guided mesh compared with conventional vaginal repair in recurrent prolapse: a randomized controlled trial. *Obstet Gynecol* 2011; 117: 242-250.
14. Detollenaere RJ, den Boon J, Kluivers KB, Vierhout ME and van Eijndhoven HW. Surgical management of pelvic organ prolapse and uterine descent in the Netherlands. *Int Urogynecol J* 2013; 24: 781-788.
15. Friedman T, Eslick GD and Dietz HP. Risk factors for prolapse recurrence: systematic review and meta-analysis. *Int Urogynecol J* 2018; 29: 13-21.
16. Espuña M, Puig M and Carmona F. De novo dyspareunia after pelvic organ prolapse surgery. *Gynecological Surgery* 2010; 7: 217-225.
17. Denman MA, Gregory WT, Boyles SH, Smith V, Edwards SR and Clark AL. Reoperation 10 years after surgically managed pelvic organ prolapse and urinary incontinence. *Am J Obstet Gynecol* 2008; 198: 555 e551-555.
18. Alas AN, Chinthakanan O, Espaillat L, Plowright L, Davila GW and Aguilar VC. De novo stress urinary incontinence after pelvic organ prolapse surgery in women without occult incontinence. *Int Urogynecol J* 2017; 28: 583-590.

19. Bugge C, Adams EJ, Gopinath D, et al. Pessaries (mechanical devices) for managing pelvic organ prolapse in women. *Cochrane Database Syst Rev* 2020; 11: CD004010.
20. van Geelen JM and Dwyer PL. Where to for pelvic organ prolapse treatment after the FDA pronouncements? A systematic review of the recent literature. *Int Urogynecol J* 2013; 24: 707-718.
21. Utomo E, Korfage IJ, Wildhagen MF, Steensma AB, Bangma CH and Blok BF. Validation of the Urogenital Distress Inventory (UDI-6) and Incontinence Impact Questionnaire (IIQ-7) in a Dutch population. *Neurourol Urodyn* 2015; 34: 24-31.
22. Cobra OK-Klapper 2012. DCHG. 2012.
23. Pollard ME, Eilber KS and Anger JT. Abdominal approaches to pelvic prolapse repairs. *Curr Opin Urol* 2013; 23: 306-311.
24. Dietz V, Schraffordt Koops SE and van der Vaart CH. Vaginal surgery for uterine descent; which options do we have? A review of the literature. *Int Urogynecol J Pelvic Floor Dysfunct* 2009; 20: 349-356.
25. de Boer TA, Milani AL, Kluivers KB, Withagen MI and Vierhout ME. The effectiveness of surgical correction of uterine prolapse: cervical amputation with uterosacral ligament plication (modified Manchester) versus vaginal hysterectomy with high uterosacral ligament plication. *Int Urogynecol J Pelvic Floor Dysfunct* 2009; 20: 1313-1319.
26. Srikrishna S, Robinson D and Cardozo L. Validation of the Patient Global Impression of Improvement (PGI-I) for urogenital prolapse. *Int Urogynecol J* 2010; 21: 523-528.
27. van Dongen H, van der Vaart H, Kluivers KB, Elzevier H, Roovers JP and Milani AL. Dutch translation and validation of the pelvic organ prolapse/incontinence sexual questionnaire-IUGA revised (PISQ-IR). *Int Urogynecol J* 2019; 30: 107-114.
28. Barber MD, Walters MD and Bump RC. Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). *Am J Obstet Gynecol* 2005; 193: 103-113.
29. Rogers RG, Rockwood TH, Constantine ML, et al. A new measure of sexual function in women with pelvic floor disorders (PFD): the Pelvic Organ Prolapse/Incontinence Sexual Questionnaire, IUGA-Revised (PISQ-IR). *Int Urogynecol J* 2013; 24: 1091-1103.
30. Constantine ML, Pauls RN, Rogers RR and Rockwood TH. Validation of a single summary score for the Prolapse/Incontinence Sexual Questionnaire-IUGA revised (PISQ-IR). *Int Urogynecol J* 2017; 28: 1901-1907.
31. European Medicines Agencym guideline *Missing data in confirmatory clinical trials*, CPMP/EWP/1776/99 Rev. 1 as amended. London.2010.

2.3 Summary of amendment

In the final version of the statistical analysis plan we clarified the method for multiple imputation. Missing primary outcome data will be imputated if results before imputation indicate non-inferiority. If non-inferiority is not shown imputation will be conducted provided the missingness and dropouts do not favor the alternative hypothesis, as stipulated in European Medicines Agency (EMA) guidance

2.3.1 Table with amendment and corresponding section

Amendment	**Corresponding section in the final version 2**
1. Clarification of multiple imputation	Section 15.6