NCT 03632447

A Prospective Randomized Efficacy Study Comparing a Pelvic Digital Health System Home Program of Pelvic Floor Muscle Exercise to Kegel Exercises in the Treatment of Stress-Predominant Urinary Incontinence

November 21, 2018

Statistical Analysis Plan

A Prospective, Randomized Efficacy Study Comparing a Pelvic Digital Health System Home Program of Pelvic Floor Muscle Exercise to Kegel Exercises in the Treatment of Stress-Predominant Urinary Incontinence (Renovia-05)

Principal Investigator: Renovia Inc.

Prepared by: Anna Modest, PhD

November 1, 2018

Table of Contents

1. Study Overview

1.2 Objectives

- Compare the efficacy of using a novel intravaginal system (leva® Plus Pelvic Digital Health System TM) to perform pelvic floor muscle exercises (PFME) compared to a Kegel exercise home program in women with stress-predominant urinary incontinence (SUI).
    - HYPOTHESIS: The leva® Plus Pelvic Digital Health System is superior to a home PFME program in the treatment of stress-predominant urinary incontinence.
    - HYPOTHESIS: The leva® Plus Pelvic Digital Health System leads to more significant improvements in pelvic floor muscle performance than a home PFME program.
- Evaluate the contribution of the Digital Health Platform in adherence and treatment of SUI for long term maintenance therapy.
    - HYPOTHESIS: The Digital Health Platform improves adherence to a long-term maintenance program of pelvic floor muscle exercises.

1.1 Design

This is a multi-center, single-blind, nested, randomized-controlled trial of 225 adult women with urinary incontinence.

For detailed design, please see available protocol. Table below from Renovia-05 protocol.

Intervention and outcomes	**Screening Visit**	**Baseline**	**4 weeks**	**8 weeks**	**6 months**	**12 months**
Eligibility and informed consent	x					
Urine Pregnancy Test (if indicated)	x					
UDI-6	x					
Medical History	x	*				
MESA questionnaire	x					
POP-Q	x[#]			x		
Brinks Scale		x*		x		
Primary Outcome Measures (PFDI, PGI-I, PGI-S)		x	x	x	x	X
PFDI, WHODAS, PISQ-IR, McGill Pain Questionnaire SESPPFE, PGI-S, PGI-I, PFIQ		x	x	x	x	X

PFM performance assessment (PFDx)		x	x	x		
3-day voiding diary	x (before baseline visit)			x		
CGI-S, CGI-I	x			x		
Complications			x	x		
Follow up questions detailing further treatments for SUI					x	X

*May be done at screening visit if < 3 months prior to baseline visit

[#] May be from a visit within 3 months of baseline visit.

2. Populations

2.1 Intent-to-Treat (ITT) Population

The ITT population is defined as all women who are randomized to either the intervention or control arm, regardless of compliance. All participants will be analyzed according to the assigned treatment arms.

2.2 Per Protocol Efficacy (PPE) Population

The PPE population is defined as all randomized participants who reported compliance with their respective treatment. Compliance is defined in section 4.6.

3. Statistical Analysis

3.1 Participant Flow

Participant flow will be described in a flow diagram. If available, the flow diagram will begin with the number of potential participants approached and then move to number of participants consented and randomized. Alternatively, if information on the number of participants approached is not available, the flow diagram will begin with the number of participants consented. The diagram will show the number of participants at each study time point, and the number lost in between each study time point with a reason for loss, if available.

3.2 Randomization

- Intervention: leva® Plus Pelvic Digital Health System™

- Control (reference): At home Kegel exercise

3.3 Primary Outcome

Efficacy of Ieva® Plus Pelvic Digital Health System: Two definitions of efficacy will be reported as two "primary" outcomes. Both will be dichotomous (yes or no). Efficacy will be assessed at the 8-week visit and will be defined as: 1, improvement of symptoms, and 2, no urinary distress.

- "Improvement" will be defined as "much better" or "very much better" on the Patient Global Impression of Improvement (PGI-I) scale. Any other answers will be considered "No improvement".
- "No urinary distress" will defined as a raw score of 2 or less on all six indicators in the Pelvic Floor Disability Index (PFDI-20) Urinary Distress Inventory (UDI-6) subscale. Raw scores of greater than 2 will be considered reports of distress.

Both will be reported as the primary outcome.

3.4 Secondary Outcomes

3.4.1 Secondary definition of efficacy

Efficacy will be defined improvement on the PGI-I *AND* no urinary distress as reported on the PFDI-20 UDI-6. "Improvement" will be defined as "much better" or "very much better" on the PGI-I. Any other answers will be considered "No improvement". "No urinary distress" will defined as a raw score of 2 or less on all six indicators in the PFDI-20 UDI-6 subscale. Raw scores of greater than 2 will be considered reports of distress.

3.4.2 Scoring of questionnaires

- **Pelvic Floor Disability Index (PFDI):** Obtain the mean value of all of the answered items within the corresponding scale (possible value 0 to 4) and then multiply by 25 to obtain the scale score (range 0 to 100). Missing items are dealt with by using the mean from answered items only. For the summary score, add the scores from the 3 scales together to obtain the summary score (range 0 to 300). This can be reported as a continuous variable or may be categorized depending on distribution of the data. For the primary outcome, "No urinary distress" will defined as a raw score of 2 or less on all six indicators in the UDI-6 subscale. Raw scores of greater than 2 will be considered reports of distress.
- **Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-IR)**: Validated scoring available from Constantine, M.L., Pauls, R.N., Rogers, R.R. et al. Int Urogynecol J (2017) 28: 1901.
- **McGill Pain Questionnaire**: The level of pain for each word can be reported as None, Mild, Moderate, Severe. Collapsing of responses is also possible (pain vs. no pain, no/mild pain vs. moderate/severe pain). Reporting of the visual analogue scale will depend on data recording. Reports of pain location will need to be manually reviewed and categorized.

- **Self-Efficacy Scale for Practicing Pelvic Floor Exercises (SESPPFE):** Sum all of the scores to calculate total score for the scale. This can be reported as a continuous variable or may be categorized depending on distribution of the data.
- **Patient Global Impression of Improvement (PGI-I) Scale:** This is only a single question and will be dichotomized as improvement ("very much better" or "much better") vs. no improvement (all other non-missing answers).
- **Patient Global Impression of Severity (PGI-S) Scale:** This is only a single question and can be reported as is.
- **World Health Organization Disability Assessment Schedule (WHODAS):** Two options are available for scoring the WHODAS. The first is a simple summary score as a continuous variable. The second option is to use the available coding provided by the WHO (http://apps.who.int/iris/bitstream/handle/10665/43974/9789241547598_eng.pdf;jsessionid=680B7F5F1FC5DE085FA4DCFD302F1176?sequence=1). A score of 17 or greater constitutes significant disability.
- **Pelvic Floor Impact Questionnaire (PFIQ)**: Obtain the mean value for all of the answered items within the corresponding scale (possible value 0 – 3) and then multiply by (100/3) to obtain the scale score (range 0-100). Missing items are dealt with by using the mean from answered items only. To obtain a summary score, add the scores from the 3 scales together (range 0-300). This can be reported as a continuous variable or may be categorized depending on distribution of the data.
- **Pelvic Organ Prolapse assessment (POP-Q)**: Assessment of pelvic organ prolapse will be conducted as medically directed. POP-Q stages may be reported as 0-IV, dichotomized as greater than or equal to Stage II vs. less than Stage II, or categorized depending on distribution of the data.
- **PFDx:** Assessment of baseline angle (rest angle), maximum angle (maximum squeeze), duration of pelvic floor lift (in seconds), valsalva angle (minimum angle), number of pelvic floor lift/relax repetitions in 15 seconds. All of these data will be reported as continuous.
- **Brink Score:** Scoring will be conducted as medically directed. All of the ratings will be summed and reported as a continuous variable.

3.4.3 Long-term treatment success

Long-term treatment success will be defined using the primary outcome definitions of improvement and no urinary distress as reported on the PFDI-20 UDI-6 at 6 months and 12 months post intervention. This can be assessed at all visits up to the 12-month assessment.

3.4.4 Short-term adherence

Short-term adherence will be defined as 80% compliance with exercises for at the 8-week visit. A more formal definition of compliance is in section 4.6.

3.5 Statistical methodology

3.5.1 General statistical analysis guidelines

The primary analysis will be conducted in the ITT population. The denominator for all data will be the total number of women randomized to each group. Two-sided p-values less than 0.05 will be considered statistically significant.

Normality of continuous variables will be assessed using the Shapiro-Wilkes test and a visual inspection of the data distribution. Continuous variables that are normally distributed will be reported as means ± standard deviation. Non-normally distributed data will be reported as median and interquartile range (25^{th} and 75^{th} percentile). Categorical variables may be collapsed as necessary depending on data distribution, although all efforts will be made to maintain pre-determined categories. Categorical variables will be presented as n (%).

Data will be examined descriptively. Crude proportions of the primary and secondary outcomes at 8 weeks will be reported. Log-binomial regression will be used to calculate risk ratios and 95% confidence intervals (CIs) to assess whether the intervention was efficacious. We will assess each portion of efficacy separately (improvement and distress) as well as combined. If differences between the intervention and control groups as seen as baseline, these may be adjusted for in a log-binomial regression.

Chi-square or Fisher's exact tests may be used to calculate p-values to assess differences between categorical variables, as appropriate based on distribution of the data. To assess differences in normally distributed continuous variables between groups, the Student's T-test will be used to calculate p-values. To assess differences in non-normally distributed continuous variables between groups, the Wilcoxon rank-sum test will be used. Log-binomial regression will be used to calculate risk ratios and 95% CIs to assess the association between the intervention and all categorical variables. Linear regression will be used to calculate difference in means and 95% CI to assess the relationship between the intervention and all continuous variables.

In the event that the log-binomial regression does not converge, modified Poisson regression with robust standard errors will be substituted. If this is consistent problem, the statistician may decide to use this approach for all calculations of risk ratios and 95% CIs.

3.5.2 Choosing potential covariates

Adjustment for potential covariates should not be necessary in the ITT approach due to the randomized study design. However, if differences in baseline characteristics are seen, an adjusted risk ratio or difference in means and 95% CI may be reported. Covariates will be considered if they differ between the intervention and control group, are related to the outcome, and are not a result of the intervention. Adjusting for covariates on the causal pathway (particularly those assessed at the 4-week visit) should be avoided. Depending on the sample size, all potential covariates may be adjusted for in the models. Alternatively, manual model selection may be undertaken, where only those covariates that change the measure of association by more than 10% are included in the model. Covariates should not be chosen only based on statistical criteria.

3.5.3 Outliers

The biostatistician will identify participants with data values that appear to be potential outliers. These values will be verified against source documents by the clinical monitors in conjunction with investigators and coordinators. Clear identification of a value as an outlier will be based on medical judgment as well as on statistical grounds. In the event that outlier values are identified, any analysis using the actual values will be followed by an analysis that reduces the outlier effect. Differences between these analyses will be discussed in the report.

3.5.4 Missing data

Analytic methods that are robust to missing data will be used. Patterns of data missingness will be described. As appropriate, missing data may be imputed or observations with missing data may be dropped.

4. Planned analyses at study time points

4.1 Testing randomization at baseline

Randomization will be assessed by comparing baseline characteristics by randomization group. Baseline characteristics will include demographics (i.e. participant age, race, other measures of SES) and medical history (i.e. gravidity, parity, history of abdominal surgery) as reported on the baseline questionnaires. In addition, the results of the initial exam assessing urinary incontinence (including voiding diary, Pelvic Organ Prolapse assessment (POP-Q), PFDx device assessment) and the results of the questionnaires administered (PFDI, PISQ-IR, McGill Pain Questionnaire, SESPPFE, PGI-S, PGI-I, WHODAS, PFIQ) will be reported for both groups. Available scoring for all validated measures will be used.

4.2 4-week assessment

The results of the PGI-S, PGI-I, PFDI, PISQ-IR, McGill Pain Questionnaire, SESPPFE, WHODAS, PFIQ, and clinical assessment of urinary incontinence (POP-Q, PFDx device assessment) will be reported.

4.3 8-week assessment – Primary outcome visit

The primary outcome, 8-week assessment, will be reported as described above. Efficacy will also be reported as both improvement on the PGI-I and no urinary distress on the PFDI-20. In addition, the results of the PFDI, PISQ-IR, McGill Pain Questionnaire, SESPPFE, WHODAS, PFIQ, PGI-S and clinical assessment of urinary incontinence (voiding diary, POP-Q, PFDx device assessment) will be reported.

4.4 Testing nested randomization

Nested randomization will be assessed by comparing baseline characteristics as well as 8-week characteristics by the three randomization groups: leva® Plus Pelvic Digital Health System group with reminders, leva® Plus Pelvic Digital Health System group with no reminders, control group. Baseline

characteristics will include demographics (i.e. participant age, race, other measures of SES) and medical history (i.e. gravidity, parity, history of abdominal surgery) as reported on the baseline questionnaires. In addition, the results of the initial exam assessing urinary incontinence (including voiding diary, Pelvic Organ Prolapse assessment (POP-Q), PFDx device assessment) and the results of the questionnaires administered (PFDI, PISQ-IR, McGill Pain Questionnaire, SESPPFE, PGI-S, PGI-I, WHODAS, PFIQ) will be reported for both groups. Available scoring for all validated measures will be used.

4.5 Long-term treatment success and outcomes at 6 months and 12 months

Long-term treatment success will be assessed using the primary outcome definitions of improvement and no urinary distress at each of the time points. The results of the PGI-S, PFDI, PISQ-IR, McGill Pain Questionnaire, SESPPFE, WHODAS, and PFIQ will be reported in the three randomization groups at 6 and 12 months. In addition, further treatment sought by participants will also be reported.

4.6 Compliance and long-term adherence

Compliance will be assessed by participant report at the 8-week study visit. Participants who report performing the exercises at least 80% of the time will be considered compliant. Exercises using the leva® Plus Pelvic Digital Health System™ should be performed twice daily, for a total of 14 times per week; therefore, participants who perform the exercise at least 11 times per week will be considered compliant. Exercises in the control (Kegel) arm should be performed three times per day, for a total of 21 times per week; therefore, participants who perform the exercises 17 times per week will be considered complaint.

In addition, for participants in the leva® Plus Pelvic Digital Health System arm, self-report will be validated using data from the device. The distribution of concordance will be reported. Allowing for some error in reporting, reports within 1 exercise per week (over or under report) will be considered concordant. This may be altered based on the data, at which point the statistician and investigators may re-define concordance.

4.7 Long-term adherence

Long-term adherence will be assessed only in the two leva® Plus Pelvic Digital Health System groups. Adherence will be defined as completing at least 80% of the exercises (at least 11 times per week) for 6 and 12 months using available data from the leva® Plus Pelvic Digital Health System device. Adherence will be compared between the leva® Plus Pelvic Digital Health System group who received reminders and the leva® Plus Pelvic Digital Health System who did not receive reminders after eight weeks.

4.8 Per protocol analysis

Distribution of baseline characteristics and primary outcome analyses will be repeated in the PPE population for the per-protocol analysis. If overall compliance in the leva® Plus Pelvic Digital Health System group is similar to compliance based on the data from the device, compliance for all groups will be defined by participant report. If concordance is poor, defined as more than 10% discordance,

additional sensitivity analyses may be necessary. Other definitions of discordance may be determined depending on the data. Given that true compliance is not known in the control group, quantitative bias analyses will be done to determine the effect of misclassification on the assessment of the primary outcome. Data of concordance in the leva® Plus Pelvic Digital Health System group will be used to determine the sensitivity and specificity of exposure misclassification in the control group.

5. Additional comments

5.1 Reporting of statistical analyses

All reporting of statistical analyses will be in accordance with the CONSORT guidelines (http://www.consort-statement.org/consort-2010).

5.2 Adverse events

Adverse outcomes should be reported at each time period. Any analyses related to adverse events will be conducted based on the distribution of the data.

5.3 Additional analyses

Additional analyses may be conducted as needed based on data at the discretion of the statistician/investigator. Study questions not addressed in this document may require additional analysis plans.

5.3.1 Correlations between PFDx and urinary distress

The PFDx measures will be compared in the women who report urinary distress on the UDI6 and women who do not report urinary distress on the UDI6. Continuous measures will be reported as described above. In addition, if UDI6 is reported as a continuous measure, Spearman or Pearson correlation, as appropriate, may be used. Future work to determine cut-off values should include sensitivity and specificity calculations and area under the curve (AUC) calculations. Additional analyses plans will be required.