**CONFIDENTIAL**

**Pelvic Floor Disorders Network**

**Protocol**

# Treatment for Mixed Urinary Incontinence: Mid-urethral Sling vs. Botox A (MUSA)

Short title: Sling vs Botox for Mixed Incontinence

**Concept Proposal:** Approved by Steering Committee: January 27, 2017
**Mini Protocol:** Approved by Steering Committee: July 19, 2019
**Protocol:** Approved by Steering Committee: December 16, 2019

**Revised:**

January 10, 2020 after DSMB review, Approved by DSMB: February 19, 2020
May 11, 2020 revised, Approved by DSMB: May 18, 2020
June 30, 2020 revised
November 9, 2020 revised, Approved by DSMB: December 18. 2020
January 22, 2021 revised, Approved by DSMB February 12, 2021

**Protocol Chair:**
Heidi S. Harvie
University of Pennsylvania
Division of Urogynecology
3400 Spruce Street, Dulles Building 5th Floor
Philadelphia, PA 19104
hharvie@pennmedicine.upenn.edu

**Protocol Committee:**
Brown/Women and Infants: Vivian Sung
Duke University Medical Center: Cindy L. Amundsen
University of Alabama at Birmingham: Holly E. Richter
Kaiser Permanente-San Diego: Shawn Menefee
University of Pennsylvania: Lily Arya (Co-Chair)
University of Pittsburgh: Christopher Chermansky
University of Texas Southwestern: David Rahn
RTI DCC: Sonia Thomas
NIH/NICHD: Donna Mazloomdoost

# ABBREVIATIONS

| | |
|---|---|
| ABC | Anticholinergic versus Botox Comparison trial |
| ATLAS | Ambulatory Treatments for Leakage Associated with Stress Incontinence trial |
| BBUSQ | Birmingham Bowel Urinary Symptom Questionnaire |
| BD | Bladder diary |
| BE- DRI | Behavior Enhances Drug Reduction of Incontinence trial |
| BPTx | Behavioral/pelvic floor therapy |
| CDF | Cumulative distribution function |
| CIC | Clean intermittent catheterization |
| CST | Cough stress test |
| DCC | Data Coordinating Center |
| DO | Detrusor overactivity |
| DSMB | Data and Safety Monitoring Board |
| ESTEEM | Effects of Surgical Treatment Enhanced with Exercise for Mixed Urinary Incontinence trial |
| EQ-5D | European Quality of Life-5 Dimensions |
| HRQOL | Health related quality of life |
| IE | Incontinence episode |
| ICI | International Consultation on Incontinence |
| ICS | International Continence Society |
| IIQ | Incontinence Impact Questionnaire |
| IRB | Institutional Review Board |
| ITT | Intention-to-treat |
| IUGA | International Urogynecological Association |
| MESA | Medical, Epidemiologic, and Social Aspects of Aging |
| MID | Minimum important difference |
| MIMOSA | Mixed Incontinence: Medical or Surgical Approach trial |
| MSM | Medical Safety Monitor |
| MUI | Mixed urinary incontinence |
| MUS | Mid-urethral sling |
| OAB | Overactive bladder |
| OAB-q | Overactive Bladder Questionnaire |
| OAB-q-SS | Overactive Bladder Questionnaire-Symptom subscale |
| OAB-SAT-q | Overactive Bladder Questionnaire-Satisfaction with Treatment Questionnaire |
| OPTIMAL | Operations and Pelvic Muscle Training in the Management of Apical Support Loss trial |
| PFD | Pelvic floor disorder |
| PFDI | Pelvic Floor Disorder Inventory |
| PFDN | Pelvic Floor Disorders Network |
| PFMT | Pelvic floor muscle training |
| PGI-I | Patient Global Impression- Improvement |
| PGI-S | Patient Global Impression-Severity |
| PISQ-IR | Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire |

| | |
|---|---|
| POPQ | Pelvic Organ Prolapse Quantification system |
| PRO | Patient reported outcomes |
| PVR | Postvoid residual |
| QoL | Quality of life |
| QUID | Questionnaire for Urinary Incontinence Diagnosis |
| RCT | Randomized controlled trial |
| ROSETTA | Refractory Overactive Bladder: Sacral Neuromodulation v. Botulinum Toxin Assessment trial |
| RUBI | Refractory idiopathic urge incontinence and botulinum A injection trial |
| SAE | Serious adverse event |
| SD | Standard deviation |
| SF-36 | Short Form 36 |
| SF-6D | Short Form 6D |
| SISTEr | Stress Incontinence Surgical Treatment Efficacy Trial |
| SUI | Stress urinary incontinence |
| TOMUS | Trial of Mid-Urethral Slings |
| TOT | Transobturator tape sling |
| TVT | Tension-free vaginal tape sling |
| TVT-O | Tension-free vaginal tape obturator |
| UDE | Urodynamic evaluation |
| UDI | Urogenital Distress Inventory |
| UI | Urinary incontinence |
| UIE | Urinary incontinence episode |
| UITN | Urinary Incontinence Treatment Network |
| UUI | Urgency urinary incontinence |
| ValUE | Value of Urodynamic Evaluation trial |
| VPFMC | Voluntary pelvic floor muscle contraction |
| 3IQ | 3 Incontinence Questions Assessment Tool |

# TABLE OF CONTENTS

**List of Figures**



**List of Tables**

## ABSTRACT

Mixed urinary incontinence (MUI), defined as the presence of both stress urinary incontinence and urgency urinary incontinence, is a challenging condition for which clinicians frequently use multiple sequential treatments that have undergone limited head to head comparison in rigorous clinical trials. **M**id-**u**rethral **S**ling vs. Botox **A** (MUSA) is a randomized 2-arm clinical trial for women who have at least moderate bother from both stress and urgency incontinence and who have failed one or more conservative treatments. The primary objective is to estimate the effect of 100 units of intradetrusor injections of Botulinum toxin A (Botox A ®) compared to mid-urethral sling surgery for the treatment of MUI in 146 women. Participants in the Botox A arm may receive one additional injection of 100 units Botox A between 3 months and 6 months after the initial injection if they have persistent urgency incontinence symptoms and meet the safety criteria. The primary outcome for this trial is the change in severity of MUI symptoms which will be measured at 6 months using the Urogenital Distress Inventory. Secondary objectives include identifying predictors of treatment failure and cost-effectiveness of treatments. Safety and adverse events data will also be collected. At the completion of this study, we will better understand whether a surgical treatment that focuses on the urgency component (Botox A) is superior to a surgical treatment that focuses on the stress component (mid-urethral sling). The trial will provide clinically useful information for two treatments that are widely used to treat MUI but for which evidence-based data are not available.

## 1. STUDY AIMS

Mixed urinary incontinence (MUI), defined as the presence of both stress urinary incontinence (SUI) and urgency urinary incontinence (UUI), is a challenging condition for which clinicians frequently use multiple sequential treatments that have undergone limited evaluation in rigorous clinical trials. The Mixed Urinary Incontinence: Mid-urethral Sling vs. Botox A (MUSA) trial will estimate the effect of intradetrusor injections of Botulinum toxin A (Botox A ®) compared to mid-urethral sling for the treatment of MUI symptoms in 146 women. MUSA is a randomized 2-arm clinical trial.

The purpose of MUSA is to:

- compare treatment with either intradetrusor injections of Botulinum toxin A (Botox A ®) or mid-urethral sling for women with MUI
- characterize patient characteristics associated with treatment response

The primary objective is to estimate the effect of intradetrusor injections Botulinum toxin A (Botox A ®) compared to mid-urethral sling for treatment of MUI in 146 women 6 months after treatment. The change in severity of MUI symptoms will be measured using the Urogenital Distress Inventory.[1]

Secondary objectives include identifying predictors of treatment failure and cost-effectiveness of treatments in this MUI population.

### 1.1. Primary Aim

To compare the effectiveness of intradetrusor injection of 100 units of Botox A to mid-urethral sling for change in MUI symptoms 6 months following treatment.

Primary Outcome: Change in severity of MUI symptoms 6 months post treatment measured using the Urogenital Distress Inventory (UDI)[1] in patients randomized to either Botox A or mid-urethral sling.

Secondary Outcomes: Change in severity of SUI and UUI symptoms 6 months post treatment measured using the stress and irritative subscales of the UDI and change in MUI symptoms 3 months post treatment measured using the UDI total score.

Null Hypothesis: There is no difference in the change in MUI symptoms between women receiving Botox A versus mid-urethral sling 6 months following treatment.

Alternative Hypothesis: Botox A is superior to mid-urethral sling for improving MUI symptoms 6 months following treatment.

### 1.2. Exploratory Aims

1. **Secondary urinary outcomes:** To compare treatment with Botox A to treatment with mid-urethral sling for improving the number of urinary incontinence episodes on bladder diary 6 months post-treatment.
2. **Predictors of poor treatment response:** To develop models to identify baseline predictors of change of MUI, OAB, and SUI outcomes measured using the UDI, between baseline and 6 months post-treatment.
3. **Quality of life and global impression:** To compare quality of life outcomes and Patient Global Impression-Improvement (PGI-I), Patient Global Impression-Severity (PGI-S) between groups randomized to Botox A versus mid-urethral sling 6 months post-treatment.
4. **Safety and additional treatments:** To describe rates of reoperation (sling revision) after mid-urethral sling and intermittent catheterization due to voiding dysfunction/partial urinary retention after Botox A detrusor injection, to compare the proportion of women in each group with UTI and recurrent UTI, rates of other serious and non-serious adverse events, and to compare the proportion of women in each group initiating additional (off protocol) treatment other than Botox A and mid-urethral sling for SUI and/or OAB.
5. **Cost-effectiveness analysis:** To determine the cost effectiveness of Botox A injection versus mid-urethral sling for the treatment of MUI symptoms on an intent-to-treat basis 6 months post-treatment.
6. **UDI MID:** To explore MIDs for UDI total score and stress and irritative subscores for this MUI population.

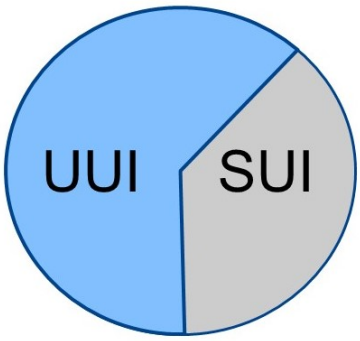
## 2. BACKGROUND INFORMATION AND SCIENTIFIC RATIONALE

### 2.1 Disease/Condition Background

Up to 50% of women with urinary incontinence have mixed urinary incontinence (MUI), a combination of stress urinary incontinence (SUI) and urgency urinary incontinence (UUI).[2-4] MUI is a frustrating problem for both patients and clinicians. Patients report that the urgency component is more bothersome than the stress component (Figure 1) and the combination of both is more bothersome than either UUI or SUI alone.[5-9] For clinicians, the treatment of MUI is challenging due to high failure rates that lead to multiple sequential

treatments that have undergone limited evaluation in clinical trials. Despite this, most clinical trials have either excluded women with MUI or focused on treating one component of the MUI defined as either "stress-predominant" or "urgency-predominant" urinary incontinence.

**Figure 1. Conceptual model of symptom bother in women with MUI**




### *2.1.1 Clinical Importance of Focusing on the Urgency Component of Mixed Urinary Incontinence*

The recently completed ESTEEM trial has provided valuable insight for the treatment of MUI, however, a key gap in knowledge remains: **what is the effect when the more bothersome UUI component is treated before the SUI component?**

The ESTEEM trial compared combined mid-urethral sling and behavioral treatment (sling-BPTx) to sling alone for the treatment of MUI in women. In ESTEEM, improvement in the total UDI score was greater in women in the combined sling-BPTx group (-128.2 points, 95% CI -146.51, -109.78) than the sling only group (-114.7, 95% CI -133.3, -96.2) but did not reach the pre-specified minimally important difference. Though authors concluded that combined sling-BPTx is not more effective than sling alone for the treatment of MUI, the study provided several additional findings that have important implications for the treatment of women with MUI.

1. **Combined sling-BPTx treatment compared to sling alone contributes to better UUI outcomes and better QOL in women with MUI.** Women in the combined treatment group reported significantly better quality of life as measured by the Incontinence Impact Questionnaire than women in the sling only group ( -131.66 vs -102.01, p = .009). The number of SUI episodes was not significantly different between groups (-1.2 vs -0.96, p = 0.2), however, women in the combined group reported significantly fewer UUI episodes than the sling only group (-1.06 vs -0.38, p = 0.016). Women in the combined group also reported fewer additional treatment with OAB medications than the sling only group (8.5% vs 17.7%, adjusted odds ratio 0.47, 95% CI 0.26, 0.85). These findings suggest that improvement in UUI and OAB outcomes were important contributors to the better quality of life in the combined treatment group and that treatments that target UUI symptoms could potentially be highly effective for the treatment of MUI. <u>However, clinical trials that have used highly effective treatments for UUI such as Botox A for the treatment of MUI are lacking.</u>
2. **Sling alone provides satisfactory treatment of SUI component and does not worsen UUI component in women with MUI**. Clinicians are frequently concerned that a mid-urethral sling in

women with MUI could potentially worsen co-existing UUI. In ESTEEM, within the sling only group, UDI-total, UDI-irritative, and UDI-stress scores at 12 months improved and exceeded the corresponding MIDs of 35 points (SD 50.4), 15 points (SD 25.6) and 8 points (SD 21.5) respectively.[10] These findings suggest that the mid-urethral sling alone is a highly effective treatment for stress component of MUI and can actually improve the UUI component. However, it remains unknown the extent to which women with MUI would require additional treatment for SUI if the more bothersome UUI symptoms were treated first.

3. **Predictors of treatment.** The ESTEEM trial showed that women with worse baseline OAB symptoms are more likely to be treatment failures at 12 months. Specifically, DO on urodynamics and previous OAB medication use were predictors of treatment failure in both groups. In addition, there was an interaction between baseline UDI-irritative score and treatment type with respect to predicting treatment failure. Women with higher baseline UDI Irritative scores randomized to sling only had higher odds of failure, compared to combined treatment. For each 10 point increase in UDI-irritative score, the adjusted odds of failure increased by 1.45 in the sling only group. This effect was not seen in the combined group.

The MUSA trial builds on the findings of ESTEEM but has several important differences from ESTEEM.

1. **The primary hypothesis of MUSA is different from ESTEEM.** The ESTEEM trial primarily treated the stress component of SUI using the mid-urethral sling and determined if behavioral treatment provided additional benefits. MUSA will determine if for women with MUI, an office-based treatment directed at the more bothersome urgency component, Botox A, is more effective than a surgical treatment directed at the stress component, mid-urethral sling.
2. **The MUI population for MUSA will be different from ESTEEM.** The ESTEEM study recruited women with MUI who desired surgical treatment for SUI symptoms. This may have allowed patients with more bother from stress symptoms to be included in the trial. The MUSA trial will recruit women with MUI who desire treatments of both the stress and urgency components. The ESTEEM trial showed that women with worse OAB are more likely to fail treatment with MUS. Therefore, in MUSA, we will recruit a patient population that will have a higher severity of urgency symptoms (objective UUI inclusion criteria of $\geq$4 UUIEs on a 3-day diary).

## 2.2 What is MUI: Current Definition and challenges

The International Urogynecological Association (IUGA)/International Continence Society (ICS) joint terminology report defines MUI as "the complaint of involuntary leakage associated with urgency and also with exertion, effort, sneezing, or coughing".[11] This definition, based on subjectively reported symptoms, has several limitations including that: 1) many women with MUI do not report clear cut SUI or UUI but simply that "they leak" 2) the definition excludes women who may have significant urgency and/or frequency without UUI; and 3) it excludes women who have detrusor overactivity in the absence of sensory urgency.

Attempts at developing more clinical meaningful definitions have not been successful. Brubaker et al of the Urinary Incontinence Treatment Network (UITN) tried to develop an operational definition of MUI based on a combination of subjective and objective criteria in a stress-predominant MUI population.[12] These criteria included 1) the frequency of SUI and/or UUI as measured by the Medical, Epidemiologic and Social Aspects of Aging (MESA),[13] 2) the presence and degree of bother for SUI and UUI as measured by the Urogenital Distress Inventory (UDI), and 3) presence of urodynamic SUI and detrusor overactivity with or without

associated leakage as measured by urodynamics. Despite testing 12 different definitions of MUI, the authors were unable to identify a definition that correlated with the clinical outcomes of the trial. Similar attempts to develop an operational definition for MUI in the urgency-predominant population of the UITN sponsored Behavior Enhances Drug Reduction of Incontinence (BE-DRI) trial also failed to identify criteria that predicted the clinical outcomes of the trial.[14]

Based on the above findings, *for the present trial, we will follow the recommendations of Brubaker et al that, until better definitions are developed, suggest that distinct descriptions of both urgency and stress subcomponents be used to characterize subjects with MUI.*

### *2.2.1 Pathophysiology of MUI: Shared Mechanisms for SUI and UUI*

While some experts argue that SUI and UUI are separate and unrelated conditions,[15] increasing evidence suggests that common pathophysiologic processes may explain co-existing SUI and UUI in women with MUI. It has long been known that anti-incontinence procedures of the bladder neck, such as colposuspension and bladder neck slings improve SUI but can worsen UUI.[16] By contrast, in the recently completed ESTEEM study, mid-urethral sling procedures alone, even without additional behavioral treatment, resulted in improvement of both SUI and UUI components of MUI. Finally, in several Botox A clinical trials, the majority of which were performed in women with urgency-predominant MUI, subjects reported improvement not only of UUI episodes but also of overall UI episodes (see Table 1). Taken together, these studies suggest that a common pathophysiologic process may contribute to co-existing symptoms of SUI and UUI as well as response to treatment.

Animal and human studies have implicated proximal urethral pathology in mechanisms of both SUI and UUI. Urethral afferent stimulation can trigger the micturition reflex in rats suggesting that in women with SUI, leakage of urine into the proximal urethra may stimulate urethral afferents and facilitate a detrusor contraction.[17] Using urodynamics, Koonings et al demonstrated urethral relaxation in 39 women with mixed urinary incontinence.[18] Combined urodynamics and ultrasound evaluations have demonstrated proximal urethral pathology, in the form of funneling and “urethral instability” (fluctuations of pressure in the proximal urethra), in women with SUI and UUI respectively.[19,20] Other shared mechanisms between SUI and UUI include cough-associated detrusor overactivity (CADO) and low maximum urethral closure pressure. In a recent retrospective review of 7009 urodynamics studies, Sinha et al reported that the prevalence of CADO, defined as detrusor overactivity immediately following the cough pressure peak on urodynamics, was 2.2% with 70% cases occurring in women with MUI.[21] The prevalence was higher (5%) in 100 women undergoing ambulatory urodynamics.[22] Lower maximal urethral closure pressure, a well-known underlying mechanism of SUI, has also been associated with UUI. In 73 women with MUI undergoing mid-urethral sling surgery, Paick et al reported that lower maximal urethral pressure at baseline was associated with a greater risk of persistent OAB symptoms following mid-urethral sling surgery,[23] suggesting that profound urethral dysfunction may contribute to OAB symptoms in women with MUI.

Translational studies also support the idea of shared mechanisms for SUI and UUI. In the RUM study by the PFDN, a variety of inflammatory and connective tissue remodeling biomarkers were measured in the urine at baseline and at 6 months following treatment with either sacral neuromodulation or Botox A in women with predominant UUI.[24] This study showed that response to Botox A was associated with significant decrease in levels of collagenase ($p < .001$) and increase in levels of matrix metalloprotease-9 ($p < .001$) and interleukin-8 ($p = .002$). Additionally, higher baseline levels of CGRP and NGF were associated with less reduction in OAB symptom bother in the Botox A group. Similar changes in inflammatory and connective tissue remodeling biomarkers were observed in women with SUI undergoing stress incontinence surgery. Chai et al

measured inflammatory and tissue remodeling biomarkers at baseline and one year following mid-urethral sling in 150 women with predominant SUI. In this study, interleukin 12p70 decreased (p = .04) and NGF increased (p =.03) one year following mid-urethral sling surgery.[25] Furthermore, change in biomarker levels was neither associated with any clinical or demographic covariates nor surgical failure.

**The shared mechanisms for SUI and UUI suggest that treatments directed at one component of MUI, either UUI or SUI, have significant potential of affecting outcomes of the other component.** In MUSA, outcomes for SUI, UUI and overall UI episodes will be measured separately using both patient reported outcome (PRO) questionnaires and bladder diary, thus providing valuable data on mechanisms that contribute to the pathophysiology of MUI.

## 2.3 Rationale for Considering Surgical Treatment for MUI

MUSA will test the effectiveness of two surgical treatments for MUI: office-based Botox A injections versus mid-urethral sling. The rationale for choosing surgical treatment over non-surgical options for MUI such as behavioral/pelvic floor therapy (BPTx) and medications is discussed below.

### *2.3.1 Behavioral/Pelvic Floor Therapy (BPTx)*

BPTx includes teaching behavioral measures designed to encourage continence and pelvic floor muscle therapy designed to strengthen and improve coordination of the pelvic floor muscles. Several studies have been performed to investigate the efficacy of BPTx for the treatment of MUI. A recent Cochrane review found that these treatments were effective for both SUI and MUI compared to placebo or no treatment, but women with pure SUI may have better outcomes.[26] Another recent systematic review concluded that BPTx, alone or in combination with other interventions, is generally more effective than pharmacologic therapies alone in treating both stress and urgency UI.[27] These studies justify recommending BPTx as a first line treatment for MUI.

Two large studies demonstrate limitations of BPTx for the treatment of more severe MUI. The UITN study BE-DRI (Burgio et al), evaluated whether combined anti-muscarinic therapy with BPTx would increase the number of women who could discontinue drug therapy (primary outcome) while sustaining a significant reduction in UUI.[28] In a population of women with pure UUI or UUI-predominant MUI, BE-DRI found that although the combination of behavioral training and drug therapy yielded improved urinary outcomes compared to drug therapy alone, the addition of behavioral therapy did not improve drug therapy discontinuation.

The recent ESTEEM trial conducted by the PFDN included 480 women with MUI who had “moderately or greatly bothersome” SUI and “moderately or greatly bothersome” UUI. In this study, women in the combined sling and BPTx group showed statistically significantly greater improvement in MUI symptoms reported on the UDI than women in the sling only group though these differences did not reach the pre-specified minimally important differences. Women in the combined treatment group reported better quality of life, less urinary incontinence episodes, and fewer additional treatments. The authors concluded that although there may be secondary benefits to adding BPTx, combined BPTx and sling was not more effective than sling alone for the treatment of MUI.

Intuitively, it makes sense to first offer non-surgical treatments such as BPTx to women with MUI. Therefore, in MUSA, women with MUI will be required to have tried and failed conservative therapy, which includes supervised behavioral therapy and/or physical therapy, before being included in the study.

### *2.3.2 Anti-Cholinergic or Beta-Mimetic Medication*

Because medical therapy is considered conservative and can be initiated in primary care offices, anti-cholinergic medications are frequently recommended before considering surgical management. However, improvement in MUI symptoms with anti-muscarinic medications over placebo is modest,[9] side effects are common, and discontinuation rate is high, ranging from 43%-83% within the first 30 days of initial prescription and only 9% at one year.[27,29] Finally, long-term anti-cholinergic use is potentially associated with potential irreversible cognitive changes.[30] The efficacy of beta-mimetic medication such mirabegron is comparable to that of anticholinergics.[31] A recent systematic review and meta-analysis of Botox A compared to medication concluded that Botox A 100 U was associated with higher odds of achieving a 100% and ≥50% decrease in urinary incontinence episodes (UIE)/day than either anti-cholinergic medications or mirabegron.[32] The combination of limited efficacy, poor compliance, and potentially serious side-effects make medication a limited treatment option for women with MUI.

**The modest success rate of BPTx in randomized clinical trials and poor adherence rate to medications point to the need to consider surgical treatment for MUI.** These findings are supported by clinical practice where women with MUI commonly opt for surgical treatment as they become dissatisfied with failure of conservative treatment and/or the need to take a medication long-term. Furthermore, women with equally bothersome UUI and SUI components commonly choose surgery, with or without a trial of conservative treatment. Thus, the "traditional" empiric treatment paradigm for MUI is associated with high failure rates and we are now challenged to consider new paradigms for MUI.

## 2.4 Rationale for Considering Botulinum Toxin A Injection for Treatment of MUI in Women

Onabotulinumtoxin A (Botox A ®, Allergan) has proved to be highly effective for the treatment of refractory UUI. Several studies on the efficacy of Botox A in improving symptoms of UUI as well as total UI episodes are summarized in **Table 1** below. Major studies are discussed below.

The ROSETTA study conducted by the PFDN was performed in women with refractory UUI and compared 190 women who received Botox A 200 units to 174 women who received sacral neuromodulation (SNS).[33] Refractory UUI was defined as a minimum of 6 urgency UIE (UUIE) on a 3-day bladder diary and persistent symptoms in spite of at least one supervised behavioral or physical therapy intervention and use or failure to tolerate a minimum of 2 anti-cholinergic medications. The primary outcome was change from baseline in mean number of daily UUIE per day over 6 months measured by monthly 3-day bladder diaries. Secondary outcomes were change in OAB symptom bother Short Form and OAB treatment Satisfaction questionnaire. At baseline, mean UUIE episodes were 5.4 (SD 2.7) in the Botox A group and 5.2 (SD 2.7) in the SNS group. Women in the Botox A group had greater reduction in mean UUIE at 6 months than the sacral neuromodulation group (-3.9 vs -3.3 episodes per day; mean difference 0.63; 95% CI, 0.13 to 1.14; p = .01). Women in the Botox A group showed significantly greater improvement in the OAB symptom bother (-46.7 vs -38.6. mean difference 8.1, p =.002) and OAB treatment satisfaction (67.7 vs 59.8, mean difference 7.8, p =.01) scores than the sacral neuromodulation group. At 24 months, no difference in mean UUIE was noted (-3.88 vs -3.5; mean difference 0.38; 95% CI -0.14, 0.89, p =.15); however, the Botox A group maintained higher satisfaction with treatment (mean difference -9.15, 95% CI -14.38, -3.9, p <.001).[34] These findings show that Botox A is highly effective for the treatment of UUI and that its beneficial effects are sustained to 24 months.

In ROSETTA, mean total UI episodes at baseline was higher than UUI episodes, being 6.0 (SD 3.0) vs 5.3 (SD 2.6) in the Botox A group. In the intent to treat population, women in the Botox A group reported

greater reduction in total UI episodes than UUI episodes, over 6 months (-4.02 vs -3.89 episodes per day). Similarly, in the clinical responder population, women in the Botox A group reported greater reduction in total UI episodes than UUI episodes, over 6 months (-4.6 vs –4.4 episodes per day). These findings suggest that the beneficial effects of Botox A may extend beyond UUIE to total UI episodes. To further explore this, we looked at ROSETTA data for a sub-group with MUI similar to what we would recruit for MUSA to provide some preliminary data of Botox A in this urgency-predominant MUI population. There were 32 ROSETTA participants randomized to Botox A with at least moderate stress bother and moderate urge bother on the Urogenital Distress Inventory Short Form (UDI-SF) questions plus 3 or more SUIE on 3-day bladder diary (average of at least 1 per day) at baseline. UDI-SF scores at baseline were worse in the "MUSA-like" group than the rest of the ROSETTA population (75.2 vs 58.2). Similarly, average daily total UIE (8.0 vs 5.7), UUIE (6.0 vs 5.5) and SUIE (1.8 vs 0.2) were higher in the "MUSA-like" group than the rest of the ROSETTA population. This supports the conceptual framework that MUI is a more severe condition than UUI alone. At 6 months, the "MUSA-like" group had a reduction of -1.1 SUIE/day (58% improvement) and -4.0 UUIE/day (66% improvement) with Botox A; the rest of the ROSETTA population had reduction of -3.4 UUIE/day (66% improvement). These findings suggest that the beneficial effects of Botox A may extend beyond UUIE to SUIE in women with MUI.
The ABC trial also conducted by the PFDN was performed in women with idiopathic UUI and compared 126 women who received Botox A 100 units to 121 women who received anticholinergic medication.[35] UUI was defined as 5 or more episodes of UUI on a 3-day bladder diary. The primary outcome was the same as in the ROSETTA study i.e. change from baseline in the mean number of UUIE over the course of 6 months on a 3-day bladder diary. At baseline, number of UUIE episodes in the Botox A group was 4.8 (SD 2.7) and 5.2 (SD 2.7) in the anticholinergic medication group. Total UI episodes were 5.8 (SD 3.1) in the Botox A group and 6.1 (SD 3.3) in the anticholinergic medication group. In this study, mean reduction in UUI episodes was similar in the Botox A and medication groups (-3.3 vs -3.4, p= 0.81). Women in the Botox A group were significantly more likely to report complete resolution of UUI (27% vs 13%, p =.003). Additionally, women in the Botox A group were significantly more likely to report complete resolution of all incontinence than the anticholinergic medication group (23% vs 11%, p =.003).

Chapple et al compared 100 units of Botox A to placebo in women with idiopathic OAB.[36] OAB was defined as ≥ 3 UUIE over 3 days and ≥ 8 micturitions per day and inadequately managed by anticholinergic medications. Co-primary outcomes were change from baseline in the number of UIE per day and proportion of patients reporting benefit on the treatment benefit scale (TBS) at week 12. At baseline, mean daily UUIE was 5.1 (SD 3.7) in the Botox A group and 5.3 (SD 3.7) in the placebo group. Baseline daily total UI episodes was 5.5 (SD 3.8) in Botox A group and 5.7 (SD 3.9) in the placebo group. Following treatment, at 12 weeks change in UUI episodes was greater in the Botox A group than the placebo group (-2.8 vs -0.82, p < .001). Change in total UI was also greater in the Botox A group than the placebo (-2.95 vs -1.03, p < .001).

Nitti 2016 analyzed the efficacy of Botox A 100 units for the treatment of refractory OAB by the number of injections received over 3.5 years.[37] Patients requested Botox injections as needed to control symptoms. The co-primary outcomes were change from baseline in the number of UI episodes per day and proportion reporting benefit on the treatment benefit scale (TBS). Mean overall change in UI episodes per day ranged from -3.1 to -3.8. Median duration of effect was 7.6 months. Long-term treatment with Botox A consistently decreased UUI symptoms with 66% to 83% patients reporting improved QOL with 1 to 4 injections.

In summary, the majority of studies investigating Botox A have been performed in subjects who have urgency-predominant MUI. These studies show that treatment with Botox A results in improvement of both

UUIE as well as total UI episodes. Though existing studies show that Botox A can improve overall UI, data on the efficacy of Botox A in a MUI population that is **not** urgency predominant incontinence is lacking.

**Table 1: Summary of Studies of Botox A for the Treatment of UUI or Urge-Predominant MUI**

| First Author | No. Pts | Inclusion Criteria | Primary Objective | Baseline UUI/UIE | FU | Change in UUI | Change in Total UI | Other Relevant Findings |
|---|---|---|---|---|---|---|---|---|
| Dmochowski 2010[38] | 313 | Refractory UUI | Dose ranging Botox 50u-300u<br>100u N=54<br>200u N=53<br>vs Placebo | 8 UUIE on a 7-day bladder diary | 12 weeks | 7-day diary UUIE reduction<br>100u -20.7<br>200u -23.0 | NA | |
| Nitti 2013[39] | 557 | UUI predom refractory | Botox 100 vs Placebo | UI/day 5.5<br>UUI/day 4.8 | 12 weeks | UUI/day not reported TBS 60.8% | UI/day -2.65<br><br>Dry 22.9%<br>≥50% 57.5% | IQOL impr 21.9 points |
| Nitti 2016[37] | 812 | Refractory UUI | Botox 100u (no. of injections)<br>1 =812<br>2 =597<br>3 =372<br>4 =264<br>12 weeks from last tx | UI 5.6 | 3.5 years | UUI/day not reported<br><br>TBS (treatment benefit scale)<br>1=74.0%<br>2=81.3%<br>3=82.1%<br>4=78.3% | UI/day<br>1 inj UI -3.3<br>2 inj UI -3.6<br>3 inj UI -3.8<br>4 inj UI -3.5 | IQOL improv<br>1 inj 66%<br>2 inj 83%<br>3 inj 76%<br>4 inj 72% |
| Chapple 2013[36] | 548 | UUI predom refractory | Botox 100 vs Placebo | UI/day 5.5<br>UUI/day 5.1 | 12 weeks | UUI/day -2.80 | UI/day -2.95 | Treatment benefit 62.8% |
| Tincello 2012[40] | 122 women | Refractory OAB/DO | Botox 200u vs Placebo | UI 6.2 | 6 months | UUI/day not reported | UI/day -4.3<br>Dry = 31% | IQOL improv 27.84 points |
| ROSETTA 2016/2018[33,34] | 381 women | Refractory UUI | Botox 200u vs SNS | UUI/day 5.4<br>UI/day 6 | 6 months<br><br><br><br>24 months | UUI/day -3.9<br><br><br><br>UUI/day -3.88 | UI/day - 4.59<br>Dry 23%<br>≥ 75% 55%<br>≥ 50% 74%<br><br>Dry 5%<br>≥75% 22%<br>≥50% 43% | Treatment satisfaction 67.7 |
| ABC 2012 [35] | 249 women | UUI predom (includes non-refractory) | Botox 100u /placebo med vs anti-cholinrtgics/ placebo inj | UUI 4.8<br>UI 5.8 | 6 months | UUI/day -3.3<br>100% UUI resolution 27%<br>≥75% UUI 54% | Dry all UI 23% | PGI-I<br>3mth 55%<br>6mth 54%<br><br>PGSC (4-5)<br>6mth 71%<br>12mth 38% |

### *2.4.1 Rationale for Dose of Botox A Injection*

Botox A is currently FDA approved for the treatment of idiopathic OAB at a dose of 100 units. The safety and efficacy of 100 units of Botox A for the treatment of UUI and urgency has been established through several large RCTs.[36-39] Since MUSA will be the first study to investigate the efficacy of Botox A in a population of women with MUI with bothersome SUI and UUI, the protocol committee made the decision that the dose used for MUSA would be to 100 units of Botox A.

### *2.4.2 Rationale for Subsequent Botox A Injection*

Botox A is a therapy requiring repeated injections. Based on product guidelines for OAB, patients should be considered for reinjection when the clinical effect of the previous injection has diminished but no sooner than 12 weeks from the prior bladder injection. MUSA will allow participants with persistent bothersome UUI to have a second Botox A injection of 100 units after 3 months, provided no contraindication to Botox A injection is present (See section 5.2.3).

## 2.5 Rationale for Considering Mid-Urethral Slings for the Treatment of MUI in Women

"Traditional teaching" is that women with MUI should not undergo anti-incontinence surgery for SUI due to the potential risk of worsening OAB symptoms. However, these recommendations are based on expert opinion and secondary analysis and retrospective studies that included traditional bladder neck slings and colposuspensions.[44-46] Most studies report that success rates of the treatment of the SUI component in a MUI population are as high as 80%.[47,48] Increasing evidence also indicates that 1) the rate of de novo UUI after mid-urethral slings is low and that 2) mid-urethral slings improve UUI and OAB symptoms in women with MUI (**Table 2**).

The recently completed ESTEEM trial was conducted in women with MUI who at baseline had 1) "moderately or severely bothersome" SUI and UUI; 2) at least one documented stress and urgency incontinence episode on the 3-day bladder diary and 3) desired surgical treatment of SUI.[49] In ESTEEM, 480 women with MUI were randomized to sling alone or sling combined with BPTx. Women in both groups had improvements in MUI symptoms as measured by the UDI-total score at 12 months (-128.15 vs -114.73, p =.03) but the difference (-13.42) did not reach the pre-specified MID of 35 points indicating that combined therapy was not more effective than MUS alone for treating MUI. However, both groups had mean within-group changes in the UDI-total, UDI-irritative, and UDI-stress scores that exceeded the corresponding MIDs of 35 points (SD 50.4), 15 points (SD 25.6) and 8 points (SD 21.5) respectively.[10] Women in both groups also reported reduction in total UI, UUI, and SUI episodes on bladder diary with significantly greater improvements in the total and urgency incontinence episodes in the combined treatment group. OAB-q symptom severity and satisfaction scores also improved in both groups and were not significantly different between groups. Finally, the rate of new or worsening urgency incontinence at or beyond the 3-month visit was 4.5% in the combined and 2% in the MUS only group. Taken together, these findings demonstrate that mid-urethral **sling alone improves overall symptoms of MUI including both the stress and urgency components and is associated with a low overall rate of worsening of de novo UUI/OAB.**

A systematic review by Jain et al in 2011 including six randomized trials and seven prospective studies reported that the overall cure rate of urgency and the UUI component of MUI after MUS was 30-85% at a follow-up of a few months to 5 years.[48] Whether authors consider MUS to be helpful or hurtful for MUI often depends on the point of view of a paper, and may also be highly dependent on the definitions used to define "persistent OAB." Some studies report that more than 50% of women with MUI experience complete resolution or improvement of OAB symptoms after MUS treatment.[50] However, other studies report that MUI is a risk factor for failure of both SUI and OAB outcomes[51] or that mid-urethral sling may exacerbate OAB symptoms. One study reported a failure rate of 42% compared to 12% for SUI outcomes in women with baseline MUI compared to those with SUI alone.[52] Whether there may be specific patient characteristics that are associated with resolution or exacerbation of OAB symptoms also remains unclear.

Two prior large trails of mid-urethral slings have been conducted in women with stress-predominant MUI. The Trial of Mid-Urethral Slings (TOMUS) trial (Richter et al) performed by the UITN, randomized 597

women with pure SUI or SUI-predominant MUI (based on MESA scores) to retropubic versus transobturator MUS.[47] Barber et al performed a second trial also comparing retropubic versus transobturator MUS for SUI.[51,53] In the Barber trial, although women with baseline detrusor overactivity were excluded, 71% had baseline UUI based on the UUI item on the PFDI-20 questionnaire.[54] In both trials, the UDI-irritative subscales improved from baseline: in the TOMUS trial, the UDI-irritative subscale improved from a mean of 41.2 (25.4) to 8.9 (15.1) at 12 months and in the Barber trial from a baseline of 3 points to 0 points at 12 months (unpublished data). The rate of de novo UUI was low in both trials (.002% and 4% respectively). However, planned secondary analysis of both trials showed that pre-operative UUI worsened post-operative outcomes. In the TOMUS trial, higher baseline MESA urge scores increased the risk of overall (objective and subjective) sling failure.[47] In the Barber trial, baseline UUI was not a risk factor for recurrent UI 1 year postoperatively, but preoperative use of anti-muscarinic medications was. However, 53% (10/19) of women taking anti-muscarinics at baseline were no longer taking them 1 year postoperatively.

In summary, the majority of clinical trials of mid-urethral sling in women with MUI have recruited women with stress-predominant MUI. While the ESTEEM trial recruited women with both bothersome SUI and UUI components, it tested a surgical intervention developed for the stress-component. These trials report high cure rate of the stress component and marked improvement and/or "cure" of OAB symptoms after mid-urethral sling; however, to date there has not been a study that has tested an intervention that primarily treats the urgency component of MUI.

**Table 2: Randomized Trials Reporting Mid-Urethral Sling Outcomes in Women with MUI*†**

| First Author | No. Pts | Inclusion Criteria | Primary Objective | % MUI at Baseline | Follow-Up | % Postop OAB and Definition | De novo OAB | Other Relevant Findings |
|---|---|---|---|---|---|---|---|---|
| Richter 2010, Richter 2017[47,55,56] | 597 | Pure SUI / SUI-predom by MESA | TVT vs TVT-O or Monarc | • 12% DO | 1 year | 10-12% *persistent* UUI (by MESA or treatment) | 0.002% New UUI | • MESA urge score risk factor for failure<br>• Baseline DO not risk factor |
| Barber[51,53] | 170 | Urodynamic SUI and no DO | TVT vs Monarc | • 71% UUI (PFDI)<br>• 14% preop anticholin | 1 year | -31% UUI postop (PFDI)<br>• 4-10% new/worse UUI (PFDI)<br>• 16% anticholin postop | 4-10% New / worse UUI | • 79% "Cure" by PGI-I <2 |
| Palva[57] | 267 | Pure SUI / SUI-predom by "detrusor instability score" | TVT vs TVT-O | • 75% frequency (UDI)<br>• 66% UUI (UDI) | 1 year & 3 year | 1 year:<br>• 22% frequency<br>• 13% UUI (UDI)<br>3 years:<br>• 36% frequency<br>• 21% UUI (UDI) | 1 year:<br>• 3-4.5%<br>3 years:<br>• 5.6-6.2% | • Only provides postop prevalence of sxs,<br>• Unclear % "persistent" or "cured" |
| Abdel-fattah[58,59] | 341 | Pure SUI / SUI-predom (undefined) | TVT-O vs ARIS | • 24% DO (N=83)<br>• 18% prior antimusc | 1 year | By BBUSQ:<br>• 23% persistent urgency<br>• 25% worsening urgency<br>• 24% persistent UUI<br>• 19% worsened | 4.3% UUI | • 52% Cure urgency<br>• 58% Cure UUI<br>• 75% "cure" by PGI-I < 2. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | UUI<br>• ~25% worsened OAB taking anticholinergics | | |
| ESTEEM Sung et al (unpublished data) | 480 | bothersome UUI and SUI, desiring surgical Rx for SUI | MUS + BPTx vs MUS only | 100% | 1 year | By UDI-irritative: Improved in both groups<br>• 45.01 vs -38.88 | • New or worsening UUI 4.5% vs 2.1%<br>• Additional treatment of LUTS 8.5% vs. 15.7% | • OABq Symptom severity and OABq satisfaction score improved in both groups |

*Excludes small, under-powered RCTs
†TVT™ (Tension free-vaginal tape, Gynecare, Ethicon Inc); TVT-O™ (Gynecare TVT™ Obturator System, Ethicon Inc); Monarc™ (American Medical Systems, Inc), ARIS® (Transobturator Sling System, Coloplast Pty Ltd)

## 2.6 Limitations of Existing Trials for the MUI Population

With the exception of ESTEEM, existing RCT data are limited because they do not focus on women with MUI and the inclusion criteria almost always require one condition to be "predominant" or "more bothersome" (e.g. UUI-predominant for most Botox A trials and SUI-predominant for most mid-urethral sling trials). Thus, women with equally bothersome symptoms are typically excluded, or may feel pressured to "choose" a most bothersome condition in order to qualify for a trial. In addition, many clinical trials focus on teasing out the effect of the intervention on urgency and stress components separately rather on improving UI as a whole. ESTEEM included women with MUI who had both bothersome SUI and UUI; however, the inclusion criteria of "desire for SUI surgery" may have resulted in the inclusion of more women with stress-predominant symptoms.

Most Botox A trials report outcomes for either UUIE or total UI episodes but not both, making it difficult to ascertain if Botox A consistently improves UIE. Similarly, many mid-urethral sling trials use a composite outcome to define failure (including response to cough stress test, negative pad test, any self-reported incontinence or incontinence on diary) and therefore it is difficult to tease out SUI and OAB outcomes separately.

## 2.7 MIMOSA Trial: First Network Trial Attempt Focused on MUI Population

In 2009 the UITN published on their experience with the "Mixed Incontinence: Medical or Surgical Approach" (MIMOSA) trial.[60] MIMOSA was designed as a pragmatic clinical trial randomizing women to nonsurgical treatment (pharmacological therapy and behavioral therapy) versus surgical treatment (mid-urethral sling including TVT, TOT, TVT-O, fascial sling and Burch). After 4-5 months of enrollment as a feasibility study, 27 women were randomized out of 1190 women screened and the study was stopped due to low enrollment. The investigators felt recruitment was challenging at least in part due to the divergent treatment approaches, but also because of the practical trial design and strict inclusion criteria.

The MUSA trial differs from the MIMOSA trial in that it compares an office-based surgical intervention (Botox A) to an operating room based surgical intervention (mid-urethral sling). Interventions being compared in MUSA are comparable to that of ROSETTA where an office-based treatment (Botox A) was compared to an operating room-based procedure, sacral neuromodulation. Therefore, the MUSA has equipoise in the treatment interventions being tested. In addition, the MUSA protocol team has carefully

selected inclusion criteria that would not be overly strict, yet still allow recruitment of a MUI population (See Section 4.2, Inclusion Criteria).

## 2.8 Summary of Known and Potential Risks and Benefits of Study Treatment

Botox A has been shown to be beneficial when used for the treatment of urgency-predominant MUI. Prior studies have documented a low rate of side-effects such as recurrent urinary tract infection and retention of urine even with a dose of 200 units of Botox A in an urge-predominant MUI population (**Tables 3 and 4**). For the MUSA population of women with MUI who receive Botox A, risks are expected to be no worse than previously reported for urgency-predominant MUI populations.

Several prior studies have documented the benefit of slings in stress-predominant MUI. The recent large ESTEEM study has shown benefit of the mid-urethral slings for both SUI and UUI in women with MUI. However, there are women with MUI who report persistent or worse UUI/OAB symptoms after mid-urethral sling and this is one potential risk, but this was <5% in ESTEEM. Risks of UTI and retention of urine following mid-urethral sling are summarized in Tables 3 and 4. For the MUSA population of women with MUI who receive mid-urethral sling alone, risks are not expected to be different compared to previous studies including MUI or SUI-predominant MUI subjects.

**Table 3: Rates of Urinary Tract Infection Following Botox A or Mid-Urethral Sling**

| | Time Point | How Defined | UTI Rate |
|---|---|---|---|
| **SLING FOR MUI** | | | |
| ESTEEM | Urine dip at baseline AND 12 months. At additional visit if clinically indicated. Negative dip required prior to baseline bladder diary. If UTI suspected at baseline, UTI was treated before collecting baseline diary. For remaining visits: dip not required solely for the purpose of bladder diary unless clinically indicated. | **Urine dip**: positive if RBC/WBC/nitrites greater than trace. **Urine culture**: performed only if clinically indicated. Positive if $\geq 10^3$ **UTI diagnosis**: based on symptoms or recent dipstick suspicious of UTI or in opinion of provider. **UTI Symptoms** include acute /recent changes such as burning with urination, acute worsening of urgency, frequency or incontinence, hesitancy of urination, blood (seen by them) in their urine, odor, lower abdominal discomfort. | 24% (sling only) vs 22% (BPTx + sling) |
| TOMUS[47] | Not in paper | Not in paper | 15.4% (TVT) vs. 9% (TOT) |
| Barber, 2008[53] | Not in paper | Not in paper | 13.6% (TVT) vs 13.4% (TOT) |
| **BOTOX A** | | | |
| ROSETTA[33] | Urine dip at baseline, every Injection visit, 2 wk, 1m, 3m, 6m | **Same as above** | 200u: 35% |
| ABC[35] | Urine dip at baseline | Not described | 100u: 33% |
| Dmochowski 2010[38] | At each FU visit and as clinically indicated | Not described | 100u: 36%<br>200u: 48% |
| Denys 2012[62] | At each FU visit and as clinically indicated | Positive dip confirmed by culture | 100u: 4.8%<br>150u: 9.1% |
| Tincello 2012[57] | At each FU visit and as clinically indicated | Not described | 200u: 30% |
| Chapple 2013[36] | At each FU visit and as clinically indicated | Urine microscopy >5/hpf + $>10^3$ cfu | 100u: 24% |
| Nitti 2016[37] | At each FU visit and as clinically indicated | Not described | 100u: 13-17% |

**Table 4: Rates of Urinary Retention Following Botox A and Mid-Urethral Sling**

| | Botox Dose | Time Point | How Defined | Retention Rate |
|---|---|---|---|---|
| **MID-URETHRAL SLING** | | | | |
| ESTEEM | NA | At baseline and 2 weeks postop | No specific PVR cutoff | 7% in both groups (sling vs sling ± BPTx)<br>Catheter after 2 weeks: 2% in both groups |
| TOMUS Richter, 2010[47] | NA | Baseline and 12 months | ISC after 6 weeks or surgery for obstructed voiding | 3.6% (TVT) vs. 0.6% (TOT) |
| Barber, 2008[53] | N/A | Baseline and 12 months | ISC after 6 weeks or surgery for obstructed voiding | 4.7% (TVT) vs. 2.6% (TOT) |
| **BOTOX A** | | | | |
| ROSETTA[33] | 200 u | At baseline and every FU visit | PVR>300 ml or PVR>200 ml and symptoms of incomplete voiding | ISC: 2wks (16%), 1 m (8%), 3m (4%), 6 m (2%), any visit (20%) |
| ABC[35] | 100 u | At baseline and every FU visit | PVR>300 ml or PVR> 150 ml and symptoms of incomplete voiding | ISC: 2wks (9%), 1m (3%), 2m (5%), 4m (3%), 6m (1%)<br><br>PVR > 150 ml: 2wks (30%), 1m (27%), 2m (20%), 4m (15%), 6m (7%) |
| Dmochowski 2010[38] | 100 u<br>200 u | At baseline and every FU visit | > 200 ml | 100u: 14.5%, 200u: 28%<br>ISC: 100u: 11%, 200u: 21% |
| Denys 2012[62] | 100 u<br>150 u | At baseline and every FU visit | > 200 ml, need for ISC | very low |
| Tincello 2012[40] | 200 u | Not specified | Need for ISC | ISC: 14-16% |
| Chapple 2013[36] | 100 u | Not specified | ISC at PVR > 200 ml | 6.9% |
| Nitti 2016[37] | 100 u | Not specified | Need for ISC | ISC 4% |

## 2.9 Significance of Proposed Study

In summary, at least four gaps of knowledge contribute to clinical challenges of treating women with MUI.

1. Though urgency urinary incontinence is the more bothersome component of MUI, most studies have focused on treating the stress incontinence-component, with or without additional treatment of the urgency component. The question that MUSA seeks to answer is: what happens to MUI symptoms if the urgency component is treated first?
2. Surgeons and researchers have spent considerable time and energy on trying to tease out stress versus urgency components of MUI to decide which component should be treated first. MUSA will provide data that will help to answer the question: Does it matter whether the urgency or the stress component of MUI is treated first with surgery?
3. Most clinical trials of Botox A have been performed in urgency-predominant MUI populations. MUSA will answer the question: What is the effect of Botox A in patients with MUI who have bothersome SUI and bothersome UUI?
4. Little is known about risk factors that predict failure of procedural treatments of MUI. MUSA will provide data to determine which women with MUI would benefit from Botox A versus which women should get mid-urethral sling.

Patients with MUI are often hopeful that their overall urinary condition will improve, but as surgeons when we offer a surgical treatment that focuses on treating the urgency component (Botox A), we cannot assure patients that it will improve their stress component. Similarly, we can provide only limited assurance that treating the SUI component with mid-urethral sling will improve their urgency component. MUSA will

provide much needed information on how outcomes can be optimized in women seeking surgical treatment of their overall urinary incontinence.

## 2.10 Innovation

This proposal is innovative for several reasons. First, the MUSA study is investigating a population of women who are often excluded from clinical intervention trials but are at high risk for treatment failure. Second, this trial will provide important information regarding baseline predictors of response such as whether urgency-predominance or stress-predominance have different optimal treatments. At the completion of this study, we will better understand whether a surgical treatment that focuses on the urgency component (Botox A) is superior to a surgical treatment that focuses on the stress component (mid-urethral sling) and will obtain predictive information that will be directly applicable to the clinical care of patients with this challenging condition.

# 3. STUDY DESIGN

## 3.1 Description of Study Design

MUSA is a multi-center randomized trial of 146 women with MUI who have elected to undergo surgical treatment for MUI. Participants will be randomized to Botox A versus mid-urethral sling.

At 6 months, the effect of treatment with Botox A or mid-urethral sling will be evaluated within a classic RCT model. The analysis will determine the effect of treatment on the primary outcome, change in Urogenital Distress Inventory (UDI) score at 6 months.

### *3.1.1 MUSA Interventions*

In MUSA, two evidence-based interventions widely used for the treatment of MUI, Botox A and mid-urethral sling, will be compared within a classic RCT setting.

The results of the MUSA trial will ultimately inform future research studies and the development of a clinical approach, potentially tailored to individual patients with MUI based on important patient characteristics.

### *3.1.2 Rationale for Not Incorporating the Type of MUI as inclusion criteria*

The protocol committee spent considerable time debating whether the type of MUI, stress- or urgency-predominant should be an inclusion criterion. Based on expert opinion, many clinicians begin the treatment of MUI with segregating symptoms of MUI into the stress and urgency components and then focusing on the most bothersome symptom (SUI vs UUI). Although many women may clearly have one condition that is more bothersome, many have equally bothersome symptoms, or cannot determine which condition is “more bothersome”. Additionally, there is little evidence that determining stress-predominance or urgency-predominance predicts response to treatment for women with MUI. The committee ultimately decided to not incorporate determining the type of MUI into the inclusion criteria as this would allow determination of whether the type of MUI is a predictor of treatment response.

### *3.1.3 Management of Patient Who is Unwilling for Randomization or Requests Treatment Off-Protocol*

We anticipate that the number of patients who are unwilling for randomization or request treatment off protocol will be low because 1) based on our eligibility criteria, subjects will already have tried and failed a variety of conservative treatments 2) every effort will be made to recruit subjects who understand the study design and are willing to undergo randomization and 3) 6 month duration of the study that would allow them to obtain the other treatment at the end of the study.

Revision of sling (e.g. for voiding dysfunction or mesh exposure) and replacement of sling that is not effective will be allowed during the study. This number was low in ESTEEM (only one subject in each arm underwent sling revision and there were no sling replacements).

Subjects who initiate additional treatment off-protocol will remain in their group assignment and primary and secondary outcome measures will be completed as scheduled. The type of additional treatment e.g. behavioral and/or pelvic floor therapy, continence pessary, medical therapy, other procedure-based treatments (posterior tibial nerve stimulation, Botox A off-protocol, mid-urethal sling off-protocol, sacral neuromodulation) will be recorded.

## 3.2 Masking

It is not feasible to mask the patients or surgeons to the intervention due to the nature of the interventions. While a placebo for Botox A is possible through saline injection during mid-urethral sling surgery, sham sling surgery would require suprapubic and vaginal incisions that would expose patients to the risks of sedation/anesthesia for the Botox A group.

Outcome assessors will be masked to treatment assignment. All post-randomization outcome measures will be assessed by masked outcome assessors. All patient-reported outcomes (PROs) will be administered prior to other clinical assessments or procedures. (**Table 5**)

**Table 5: Masking in MUSA**

| Study Individual | Masking |
|---|---|
| Study participant | No |
| Study surgeon | No |
| Outcome assessors | Yes |

## 3.3 Randomization and Stratification

Patients will be assigned to one of the two treatment groups with a randomization sequence prepared and maintained centrally by the Data Coordinating Center (DCC). Allocation to the treatment groups will be 1:1. Randomly ordered permuted blocks will be used, with block sizes known only by the DCC. The web-based data management system will provide the treatment assignment for each participant as she is randomized. Thus, the allocation sequence will be concealed from clinical site staff.

Randomization will be stratified by clinical site and age group ($\geq 65$, $< 65$).

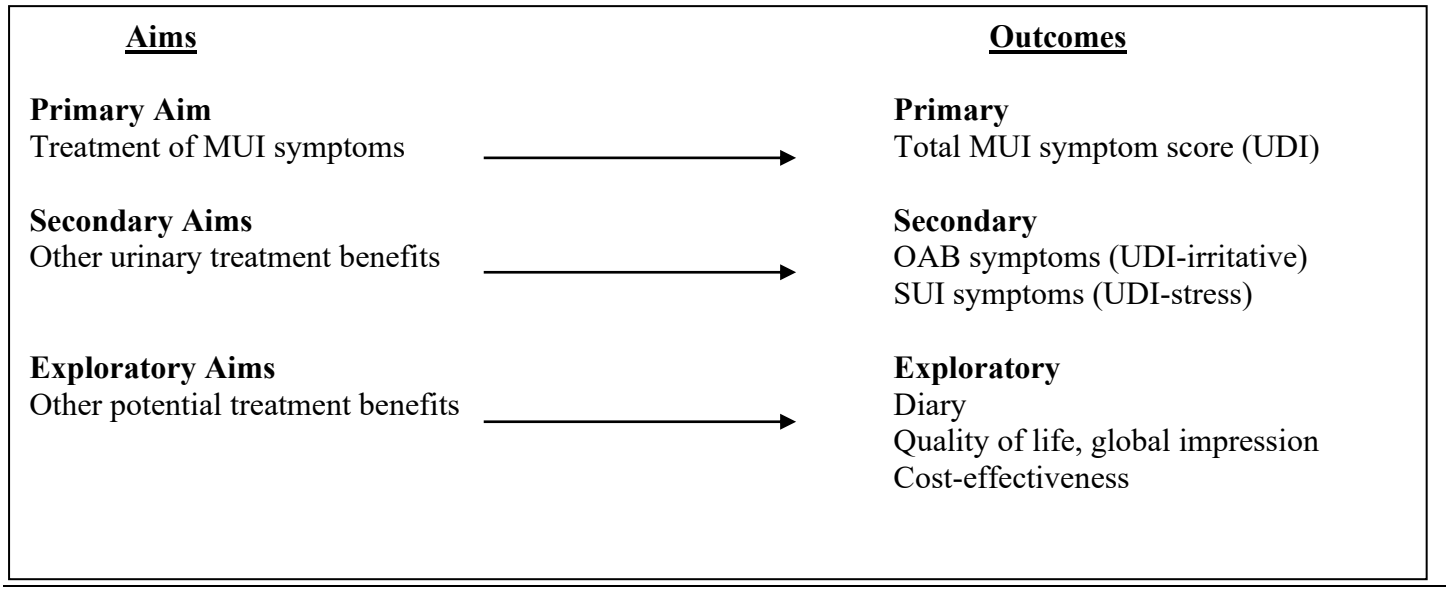
Age is a well-known predictor for reduced response to treatment for both Botox A and sub-urethral sling.[64-66] In women with refractory UUI undergoing Botox A injection in the ROSETTA study, age ≥ 65 years was associated with 3.3 fold reduced odds of ≥ 75% resolution of UUI episodes on bladder dairy and significantly less improvement on the OABq-SF than women < 65 years of age.[65] Increasing patient age was also associated with decreased odds of satisfaction after sling surgery for stress urinary incontinence.[66] Therefore, women will also be stratified by age < 65 years or ≥ 65 years.

Prior studies suggest that preoperative severity of UUI can affect treatment outcomes for UI. The "severity" of UUI has been identified as a potential risk factor for treatment failure for mid-urethral sling.[67,68] Conversely, in the ROSETTA trial, greater number of pretreatment UUI episodes was associated with greater response to treatment with Botox A injection.[64] In ESTEEM, subjects with MUI were stratified based on the number of UUI episodes. Subjects in ESTEEM were not stratified based on the number of SUI episodes because both groups were receiving the same treatment for SUI (mid-urethral sling). Similarly, in TOMUS,[47] there was no stratification on SUI severity because both groups were receiving a mid-urethral sling (retropubic mid-urethral sling versus transobturator mid-urethral sling).

The protocol committee had extensive discussions on whether subjects in MUSA should also be stratified based on the severity of UI. The definition of severity of UI for the MUI population is complex. What is important: overall UI or separate UUI and SUI severity measures? How do the various measures of UI severity differentially affect treatment failure with Botox A versus mid-urethral sling? Is there interaction between the level of UUI and SUI severity? There is limited data to inform these questions. One possible approach would be to stratify based on both UUI and SUI severity; however, this would create significant complexity in the randomization process and analysis plan. Given that subjects in MUSA will have moderate to severe UUI **and** SUI symptoms, have failed conservative treatments, and will be able to receive highly successful treatments for both UUI (Botox A) and SUI (mid-urethral sling), the team agreed that women will not be stratified based on incontinence severity.

## 3.4 Outcomes

**Figure 2. MUSA Outcomes**



### *3.4.1 Detailed Description of Primary Study Outcome*

The primary outcome for this study is the mean change from baseline in UDI-total score at 6 months after the intervention is administered (mid-urethral sling or Botox A). Outcomes at 6 months will assess the effectiveness of the study treatments.

We have elected to use the total UDI as the primary outcome because it is a validated, condition-specific patient-reported outcome (PRO) measure. A PRO is a measurement of any aspect of a patient's health status that comes directly from the patient (without interpretation by the physician, researcher, other). In clinical trials, symptom indices and quality of life PRO instruments are increasingly being used as primary outcomes and supported by federal agencies.[69,70]

The long form of the UDI is a 19 item, validated UI symptom specific PRO measure with 3 subscales: stress, irritative, and obstructive symptoms.[1] The UDI fulfills the minimum qualities needed for interpretation in a clinical trial including construct, criterion and discriminant validity. Construct validity (convergent) was originally established by demonstrating significant correlation between the overall UDI and its subscale scores with the number of IEs on 7-day diary and pad tests. Criterion validity was established by correlating total and subscale scores with physician diagnoses. The UDI can effectively discriminate between known UI clinical groups and diagnoses (specifically genuine SUI, urodynamic detrusor overactivity, or mixed) and is responsive to change. Higher scores represent more severe disease or bother from the patient perspective.

The UDI will be the primary outcome because it has all the characteristics that are important for a MUI population:

1. It captures a meaningful outcome from the patient perspective, incorporating both the presence and bother of SUI and OAB symptoms.
2. The overall UDI score includes both a stress and irritative subscale, allowing us to comprehensively capture both SUI and OAB symptom outcomes.
3. It includes 3 UI items that are not necessarily specific to stress or urgency and thus can help capture UI episodes for which patients cannot clearly distinguish as SUI or UUI.
4. It can capture both improvement and worsening of preexisting symptoms, but also the development of new urinary symptoms.[56]

Another important advantage of using the UDI is a primary outcome measure is that its minimum important difference (MID) has been extensively studied. From the patient perspective, MID represents "the smallest difference in score in the domain of interest which patients perceive as beneficial..."[28] Statistically, *the minimum important difference (MID) of a measure is a score change that reflects a clinically meaningful response to treatment and represents the "between group" criterion that needs to be met or exceeded in order for study results to be considered clinically meaningful.* From the clinical trial perspective, the MID represents the magnitude of benefit for which trials should be powered to minimize type 1 and type 2 errors.

**Table 6** shows the MID of the UDI from three large clinical trials. Prior estimates of the MID for the UDI were obtained in urge-predominant and stress-predominant MUI populations by Dyer et al and Barber et al respectively. The most recent estimate of MID of the UDI is from the ESTEEM trial which recruited subjects that are most similar to the MUSA population: women with MUI who have bothersome stress and bothersome urgency incontinence. The baseline mean UDI score of ESTEEM is higher than that of the BE-

DRI and ATLAS populations. Using the anchor-based method, ESTEEM recommends a MID of -26 for the UDI total, -10.2 for the UDI-irritative subscale, and -5.4 for the UDI-stress subscale.

**Table 6: UDI MID**

| Trial / Author | Population | Endpoint / Intervention | UDI Component | Anchor-Based MID | Distribution-Based MID (1/2 SD) | Recommend MID | SD |
|---|---|---|---|---|---|---|---|
| ESTEEM | MUI<br>Baseline mean UDI Score: 178 ± 43, UDI-irritative 66 ± 20, UDI stress 86 ± 18 | Sling vs. Sling + Behavioral at 3, 6, 12 months | UDI-total | -26.1 | | -26 | 43 (at bl)<br>55 (change)<br>46 (chg model) |
| | | | UDI-irritative | -10.2 | | -10 | 20 (at bl)<br>25(change)<br>21 (chg model) |
| | | | UDI-stress | -5.4 | | -5.4 | 18 (at bl)<br>28 (change)<br>23 (chg model) |
| BE-DRI, Dyer 2011 [71] | Pure urge/Urge-predominant MUI.<br><br>Baseline mean UDI score 120 ± 49, UDI irritative 58 ± 22 | 8 month Meds +/- BPTx | UDI-total | -45 to -35 | -25 | -35 | 50 |
| | | | UDI-irritative | -20 to -15 | -11 | -15 | 26 |
| ATLAS, Barber 2009 [10] | Pure SUI/SUI predominant MUI<br><br>Baseline mean UDI score 80 ± 40, UDI stress 47 ± 19 | 3 month Pessary vs BPTx vs both | UDI-total | -22.6 to -6.4 | -21.9 to -18.8 | -11 | 36 |
| | | | UDI-stress | -16.5 to -4.6 | -10.6 to -9.1 | -8 | 21 |

There are at least 3 different ways to analyze the UDI scores:

#1. Compare postoperative *mean UDI scores* between groups at 3 and 6 months

#2. Compare *mean changes* (delta) in UDI scores from baseline to 3 and 6 months between groups *(preferred, see below)*

#3. Dichotomize "success" and "failure" as women who achieved a 26-point improvement versus those who did not (also known as "responder analysis")

For analysis, we plan to compare *mean changes* (delta) in UDI scores from baseline to 3 and 6 months between groups (option #2). We prefer this to comparing postoperative mean UDI scores between groups (option #1) because 1) it will account for baseline differences in UDI scores between groups, which will be important in the unlikely event that baseline scores between groups differ in spite of randomization and 2) UDI scores typically have a distribution that is highly skewed, but differences from baseline are more likely to be normally distributed. We have elected to not dichotomize patients into responders and non-responders based on MID because information such as "slightly improved" is typically lost when lumping subjects into groups and furthermore, patients may worsen significantly on one of the subscales and yet be classified as a "success".

### 3.4.1.a Advantages of the UDI Over Other Outcome Measures

The team discussed using bladder diary, IIQ (Incontinence Impact Questionnaire), and global impression of severity or improvement. Arguments against each of these were based on the following rationale:

1. Problems with using bladder diary as primary outcome:

    a. Diary does not capture a meaningful patient outcome- It is becoming increasingly clear that typical clinical trial endpoints such as reduction in the number of incontinence episodes (IEs), voided volumes, etc. do not capture what is meaningful to patients. For example, having 3 large urinary urgency leaks a day may be more bothersome than having 20 small stress leaks or having 20 urgency associated voids may be more bothersome than having 1 UUI episode. Diary IEs also have limited correlation with patient satisfaction.[72] Finally, bladder diaries have been shown to be less reliable in women with MUI, particularly for the SUI component.[73]

    b. Diary cutoffs to define improvement for MUI are unknown-What percent improvement for the SUI component and for the UUI component is clinically important for a woman with MUI? Any cutoffs chosen would be arbitrary.

    c. Using IEs on bladder diary as a primary outcome would require a minimum number of IEs (approximately 3-4 IE/3days) at baseline to be able to detect a change. The protocol team felt that setting such strict inclusion criteria would be too limiting to allow recruitment of a good range of MUI severity (see Inclusion Criteria, Section 4.2).

For all of these reasons, the team decided against using bladder diary IEs as the primary outcome and to instead focus on measures that can capture outcomes from the patient perspective. However, IEs on the bladder diary will be a secondary outcome because this outcome measure has been widely used in other urinary incontinence trials.

2. Problems with using a quality of life measure as primary outcome

    The protocol committee had extensive discussion about using a quality of life measure, such as the Incontinence Impact Questionnaire (IIQ) as the primary outcome. The IIQ long form measures the impact of UI on the QOL and is therefore suitable for capturing the effect of treatment of either stress or urgency component on QOL.[1] Ultimately, the team decided to not use the IIQ as a primary outcome because the IIQ does not allow measurement of the impact of the interventions individually on the stress and urgency components. Though there are some women with MUI who may not be able to clearly distinguish all UI episodes as stress- or urge- related, the team felt it would be important to capture the effect of the interventions individually on symptoms. Finally, the MID of the IIQ was determined in stress-predominant and urge-predominant populations and has less validity for a MUI population.[10,71]

3. Problems with using global impression measure as primary outcome

    Several Botox A trials have used a patient's overall global impression of improvement as a primary outcome measure (see table 1). Although this outcome is simple and captures a clinically meaningful outcome, for our trial it could potentially introduce bias. Because it is not feasible to mask subjects in MUSA to the intervention, a single, subjective global impression item would be subject to bias. For example, if subjects in the Botox A group were more likely to ask for additional treatment and report they were not "improved" because they knew there was another potential surgical treatment available that they did not receive, this would bias our study towards a higher failure rate in the Botox A group (making the mid-urethral sling seem more effective than it really is). The challenges of masking or designing a sham procedure for the control group for MUSA have already been noted above (Section on Masking 3.2).

### 3.4.1.b Rationale for Timing of Primary Outcome

In framing this question, the protocol committee considered the time-point at which a difference in outcome would lead to recommendation of Botox A versus mid-urethral sling as treatment for MUI in clinical practice. While three month outcomes are common for Botox A studies (table 1), longer-term outcomes of 1 year or more are "standard" for surgical trials with mid-urethral sling (table 2). The choice of 6 months for the primary outcome has precedent in our patient population for both Botox A and mid-urethral sling. For Botox A, both the ABC[35] and ROSETTA[33] trials had 6 month primary outcomes. For mid-urethral sling, while ESTEEM's primary outcome was 12 months, results for UDI total score, UDI stress score, UDI-irritative score, total UIE/day, UUIE/day, and SUIE/day did not change from 6 to 12 months within the sling only group (**Table 7**). In MUSA, the 6 month outcome will capture the short-term effect of therapy with either Botox A or mid-urethral sling for MUI before additional treatments are allowed.

The time of procedure (mid-urethral sling surgery or Botox A injection) will serve as time = 0. If a participant is randomized but procedure is never performed, then Time 0 will be the planned surgery date.

**Table 7: ESTEEM outcomes at 6 and 12 months in the sling-only arm**

| Measure (Sling arm) | 6 month mean change | 12 month mean change |
|---|---|---|
| UDI total score | -118 | -114 |
| UDI stress score | -64 | -61 |
| UDI irritative score | -40 | -39 |
| UUIE per day | -0.57 | -0.38 |
| SUIE per day | -0.96 | -0.95 |

### 3.4.1.c Management of Subjects Who Request Other Non-Study Treatment for SUI and/or OAB After Botox A or Mid-Urethral Sling

The overarching goal of MUSA is to evaluate the effect of Botox A or mid-urethral sling on improving both SUI and UUI outcomes in women with MUI. Therefore, other non-study treatment for urinary incontinence (either SUI or UUI) should not be offered during the 6 month study period. We anticipate that the number of patients who request treatments other than Botox A or mid-urethral sling will be low (See section *Management of patient who are unwilling for randomization or request treatment off-protocol).* Patients requesting other non-study treatment during the 6-month study period will be asked to take advantage of treatments offered per-protocol and defer other treatments until after 6 months. Any non-study treatment will be recorded as a protocol deviation.

### *3.4.2 Secondary Outcomes*

The UDI total score at 3 months is a secondary outcome. Because MUI includes both SUI and UUI, it is important to be able to report SUI, UUI and other OAB symptom outcomes separately. The UDI-stress

subscale and UDI-irritative subscale will be important secondary outcomes for which MUSA will be powered to detect differences and each will have a priori analysis plans (see Section 6, Statistical Considerations).

### 3.4.2.a Urgency Urinary Incontinence/Overactive Bladder Symptom Outcomes

Unlike ESTEEM, "desire for surgical treatment of SUI" is not an inclusion criterion for MUSA and there is an objective UUI inclusion criteria of ≥4 UUIEs on a 3 day diary. Therefore, it is likely that the MUSA population may have more prominent UUI symptoms than ESTEEM. A potential important clinical problem in this population is the potential for persistent or worsening OAB symptoms after mid-urethral sling, it is highly important to capture and report on the cardinal symptoms of OAB (urgency, frequency, urgency urinary incontinence and nocturia) from the patient perspective. The UDI-irritative subscale measures symptom burden, impact, and changes related to OAB which are important aspects that cannot be directly observed or otherwise measured. It is highly responsive to treatment-related change and is able to discriminate among levels of change in all bladder diary variables (urinary urgency, frequency and urge incontinence) and patient ratings of treatment benefit. Particularly for MUSA, this comprehensive OAB measure will be important to understanding how mid-urethral sling may affect all OAB symptoms individually and as a whole.

### 3.4.2.b Stress Urinary Incontinence Symptom Outcomes

Botox A has been shown to reduce overall UIEs in a urge-predominant MUI population (see section 2.4). Additionally, the common presence of cough-induced DO suggest that Botox A could potentially improve SUI symptoms in patients with MUI. However, direct data on whether Botox A can improve SUI symptoms are lacking. Therefore, it is important that SUI outcomes are reported separately in MUSA.

Prior studies have also reported higher failure rates for SUI in women with MUI. In ESTEEM, improvement in stress score of the UDI was greater in the combined sling-BPTx group than the sling only group though these differences were not statistically significant (-67.11 vs – 61.65, p =.08). The MUSA population is different from ESTEEM in that they will already have "failed" behavioral therapy. Therefore, they may be at higher risk for persistent SUI symptoms. Other studies have also reported higher failure rates for SUI for women with MUI undergoing sling surgery. One study by Paick et al evaluated 274 women, of which 73 had MUI and reported cure rates for SUI to be 78% for the MUI group and 95% for the pure SUI group.[23] They also reported that low maximal urethral pressure at baseline was associated with a greater risk of persistent OAB, suggesting the possibility that profound urethral dysfunction may contribute to persistent symptoms. In ESTEEM, baseline urodynamic evaluation (UDE) parameters that predicted failure included detrusor overactivity, lower volume at first urge, and Valsalva leak point pressure <60 cmH2O. A study by Gleason et al using data from the University of Alabama at Birmingham including 534 women with MUS found that a lower proportion of women with MUI had SUI success compared with the SUI only group (64% vs 84.5%, p < 0.001). [74] Therefore, as a secondary outcome, we will also compare SUI outcomes between women randomized to Botox A versus mid-urethral sling. SUI symptoms will be measured using the UDI-stress subscale.

### *3.4.3 Other Outcomes*

#### 3.4.3.a Other UUI/OAB Outcomes

1. Bladder diary – Change in IE frequency and type, number of urgency incontinence episodes, urgency incontinence severity with voids, number of diurnal voids, and number of nocturnal voids will be assessed and compared between groups at 3 and 6 months.
2. Overactive Bladder treatment satisfaction (OAB-SAT-q)[75] The OAB-SAT-q is an 11 item instrument designed to assess patient satisfaction with treatment in a clinical setting. There are three 3-item subscales (Satisfaction, Side Effects, Endorsement) and two single items (Convenience, Preference). Response options are presented on 4-, 5-, and 6-point Likert scales. It has demonstrated good psychometric properties in OAB/UUI patients receiving treatment. We will compare OAB-SAT-q scores at 3 and 6 months between treatment groups.
3. Overactive Bladder Questionnaire- subscales (OAB-q) – The OAB-q is a validated, responsive questionnaire that includes 8 symptom bother items (SS) and 25 health related quality of life (HRQOL) items of 4 subscales (coping, concern, sleep, and social interaction).[76] In a systematic review of UI questionnaires by Avery et al, the OAB-q was rated as "grade A", highest recommendation specifically for OAB symptoms.[77] Each item is rated on a 6-point Likert scale ranging from "not at all bothered" to "a very great deal bothered" for symptom items and "none of the time" to "all of the time" for HRQOL items. Subscales are summed and transformed into scores ranging from 0-100 with higher bother scores indicating increasing symptom bother and higher HRQOL scores indicating better quality of life.[78,79] We will compare change from baseline in OAB-q subscale scores at 3 and 6 months between treatment groups.

#### 3.4.3.b Quality of Life/Global Impression

We will compare change from baseline score in the measures below at 3 and 6 months between treatment groups.

1. Incontinence Impact Questionnaire Long Form (IIQ-LF)[57]
2. Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-IR)[80]
3. European Quality of Life-5 Dimensions (EQ-5D) [84]
4. Short Form 36 (SF-36)[81] and Short Form 6D (SG-6D)[82]
5. Patient Global Impression of Improvement (PGI-I) and Patient Global Impression of Severity (PGI-S).[83] Observed PGI-I at 3 and 6 months will be compared between groups.

#### 3.4.3.c Safety/Additional Treatment

1. Additional re-treatments for SUI or UUI within 6 months of treatment, and type of re-treatment
2. Return to OR for sling revision due to worsened OAB symptoms, urinary retention, mesh exposure, repeat sling
3. Rate of urinary retention, UTIs and recurrent UTIs
4. Rate of other serious and non-serious adverse events.

### 3.4.3.d Cost-Effectiveness Outcomes

The cost-effectiveness analysis will be conducted from the health care sector and societal perspectives and will be expressed as incremental cost required to produce one additional unit of quality-adjusted life year (QALY). Data on each subject's use of medical and non-medical resources, related to urinary incontinence will be collected during the follow up period. Direct and indirect costs of the treatment of urinary incontinence with Botox A compared to mid-urethral sling (MUS) and women's preference for health states for improvement in urinary incontinence will be estimated.

We plan to capture incremental direct health care, direct non-medical, and indirect resource use related to study interventions and complications and other urinary incontinence management (such as other UI treatment, UI products and management of side effects). Costs will be estimated using the resource costing method. Direct medical service use collected from each study case report form and direct non-medical and indirect costs collected from patient questionnaires are monetized by multiplying the number of units of each resource use by the average unit cost of this item in dollars. Detailed individual cost data will not be collected. This method allows a consistent capture of resource use when costs are incurred across multiple health systems or payers. Detailed case report forms, that include the interventions performed (e.g. mid-urethral sling surgery and Botox injections) and clinical events (e.g. complications and additional treatment) will be completed by the study coordinator at study visits. Patient questionnaire on direct non-medical costs (e.g. pads, laundry) and indirect costs (e.g. time, lost productivity) will be completed at study visits 2 weeks and 6 months. Data from medical resource types (physician visits, medications, hospital admissions and emergency room visits) will be collected. Cost for each direct medical service use, direct non-medical items, and indirect items will be assigned based on national Medicare reimbursement rates or other standardized national unit costs as indicated in the following **Table 8**.

**Table 8: Resource Utilization Data Collection and Price Data Source by Utilization Category**

| Service | Source Documentation | Price Weight |
|---|---|---|
| Surgery: mid-urethral sling | Case Report Form | Medicare reimbursement |
| Botox injection | Case Report Form | Medicare reimbursement |
| Medication | Case Report Form | Drug Red Book |
| Physician visit | Case Report Form | Medicare reimbursement |
| Complication: surgery | Case Report Form | Medicare reimbursement |
| Complication: hospitalization | Case Report Form | Medicare reimbursement |
| Complication: ER visit | Case Report Form | Medicare reimbursement |
| CIC products | Case Report Form | Average national cost |
| UI products | Questionnaire | Average national cost |
| UI laundry / dry cleaning | Questionnaire | Average national cost |
| Time | Questionnaire | Average national cost |
| Lost Productivity | Questionnaire | Average national cost |

*Rationale for Using the EQ-5D and SF-6D to Measure Utility Values*

The European Quality of Life-5 Dimensions (EQ-5D) (EuroQol Group, http://www.euroqol.org), preference-based utility index algorithm will be used to calculate each subject's utility score.[84] This instrument will be collected at baseline and follow up study visits (3 months and 6 months). The EQ-5D has 5 attributes (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) with 3 levels each for a possible 243 unique health states. The EQ-5D scoring function is based on the time-tradeoff method with UK Scores ranging from -0.59 to1.00 and US Scores from -0.11 to 1.00. This instrument has been previously validated in women with urinary incontinence and used in women with urinary incontinence, including ESTEEM.[41,42,62,63] These data will be used to compare change in QALYs between the two treatment groups. We are choosing to use a general scale to calculate change in utilities (rather than condition-specific) to allow for comparison of cost-effectiveness results with other interventions and diseases.

The Short Form 6-Dimensions (SF-6D) is a single summary preference based or utility measure of health derived from 11 items of the SF36 and will also be used to calculate each subjects utility score.[81,82] This instrument will be collected at baseline and follow up study visits (3 months and 6 months). The SF-6D has 6 dimensions (physical functioning, role limitation, social functioning, pain, mental health, vitality); each dimension has 2 to 6 levels for 18,000 possible unique health states. The SF-6D scoring function is based on the standard gamble method. The SF-6D preference-based measure is a continuous outcome scored on a 0.29 to 1.00 scale, with 1.00 indicating "full health. This instrument has been previously validated and used in women with urinary incontinence and pelvic organ prolapse.[85-87] We are using both the SF-6D and the EQ-5D to measure utility scores because though both the SF-6D and EQ-5D weigh health states on a scale of 0 (dead) to 1 (optimal health) and the MID values for both instruments are similar, they are not directly comparable. Studies have demonstrated that the SF-6D does not appear to describe health states at the lower end of the scale as well as the EQ-5D but is better able to describe health states and detect improvements towards the top of the utility scale.[88]

Further data is needed to determine which instrument performs better as a preference-based utility and general HRQOL measure for women with MUI or if both are acceptable.

A questionnaire to measure direct non-medical costs (e.g. pads, laundry) and indirect costs (e.g. time, lost productivity) will be administered. Based on similar questionnaires used in SISTEr,[89] ValUE[90] and ESTEEM[49] studies, this instrument should take approximately 15 minutes for a subject to complete at 3 months and 6 months.

## 4. SELECTION OF PARTICIPANTS

Adult women aged 21 or older with bothersome MUI (defined as bothersome SUI and UUI) will be eligible.

### 4.1 Eligibility Criteria/Rationale for Inclusion/Exclusion

#### *4.1.1 Defining the MUSA MUI Population*

The MUSA population consists of women with bothersome MUI. Women in MUSA must report subjectively bothersome SUI and UUI and must also objectively demonstrate SUI and UUI. In determining eligibility criteria, the protocol committee considered the ultimate goal of MUSA. The population of MUSA are women with MUI who would benefit from either Botox A or mid-urethral sling but it is unclear whether treatment should begin with Botox A or mid-urethral sling. Therefore, participants must have bothersome

symptoms, however, criteria are not so strict that women on either the moderate or severe end of the spectrum are excluded or recruitment is hindered (as was in MIMOSA).

A single instrument that provides objective documentation of both SUI and UUI does not exist. A bladder diary allows documentation of both conditions; however, women are frequently unable to clearly characterize a leakage episode as SUI or UUI. Similarly, more invasive UDE testing also allows objective documentation of SUI and detrusor overactivity (DO) but DO correlates poorly with OAB symptoms and UDE does not predict treatment outcomes of either SUI or UUI.[90-93] Therefore, similar to ESTEEM, MUSA will use a variety of instruments to identify women with MUI who have both subjective and objective evidence of MUI. This includes subjective documentation of at least moderately bothersome SUI and UUI on a PRO (UDI), objective documentation of UUI on bladder diary, and objective documentation of SUI by CST or UDE.

*Rationale for Bladder diary criteria for MUSA*

In determining the number of episodes required on bladder diary for inclusion into MUSA, the protocol committee reviewed bladder diary criteria for existing SUI and UUI trials (see tables 1 and 2 on Botox and mid-urethral sling). SUI trials typically have not used a bladder diary inclusion criterion; objective criteria of SUI is met by CST or UDE. Most Botox A trials have specified the number of UUI episodes in their inclusion criteria. The ESTEEM trial has shown that worse OAB (DO on urodynamics, previous OAB medication use, and for the sling only group, higher baseline UDI irritative subscore) is a predictor of failure to treatment, based on combined subjective (not meeting MCID for UDI-total score) and objective (not achieving >70% improvement on mean IE/day or additional SUI / UUI treatment) criteria . The team decided that at least 4 UUI episodes on a 3-day diary would be appropriate for documenting objective UUI and allow recruitment of a MUI population with higher UUI bother than ESTEEM while not necessarily being UUI-predominant. In ESTEEM, 70% of participants at baseline had at least 4 UUIE in 3 days (≥1.3 UUIE/day on average); this population had a mean 3.7 UUIE/day, 2.5 SUIE/day, 6.6 total UIE/day, 70 UDI-irritative score, and 85 UDI-stress score and an overall 20% failure rate at 3 months. The number of UUI incontinence episodes has purposefully not been set as high as previous Botox trials (Table 1) because the MUSA population is not limited to a refractory UUI or UUI-predominant population and to allow recruitment of a broad spectrum of women with MUI. Additionally, MUSA's primary outcome is not defined by diary improvement.

*Rationale for Urodynamic evaluation inclusion criteria for MUSA*

The protocol committee considered the utility of including UDE as a MUSA inclusion criteria. Given that this is the first trial to compare treatment with Botox A versus mid-urethral sling in women with MUI, baseline UDE information was considered important. In addition, other studies in women with MUI undergoing mid-urethral sling surgery have reported association of baseline UDE parameters with treatment failure. Paick et al reported that low maximal urethral pressure at baseline UDE was associated with a greater risk of persistent OAB after mid-urethral sling[23]. In ESTEEM, baseline UDE parameters that predicted failure included detrusor overactivity, lower volume at first urge, and Valsalva leak point pressure <60 cmH2O. Thus, standardized UDE will be performed before study interventions.

There are no specific UDE parameters that determine eligibility for this trial beyond diagnosis of SUI defined by a positive CST or UDE. Patients in MUSA will undergo UDE testing, primarily to allow evaluation of variables that may predict clinical outcome. If patients have already undergone UDE prior to enrollment, results performed within 18 months of enrollment will be allowed if studies meet MUSA

standards. MUSA will use UDE standards similar to ESTEEM. For those women who have not had UDE in the past 18 months, UDE will be performed before randomization.

### *4.1.2 Targeting a Population That is Distinct From ESTEEM*

While several inclusion criteria of MUSA are similar to that of the recently completed ESTEEM study, there are also important differences. First, inclusion criteria for the number of UUIE per day is higher in MUSA than ESTEEM, thus ensuring recruitment of a population with higher UUI bother (see above *Defining the MUSA MUI Population).* Second, unlike ESTEEM, desire for surgical treatment for SUI symptoms is not an inclusion criterion for MUSA. Removing this criterion ensures that participants are eligible for both Botox A and mid-urethral sling. Additionally, removing this criterion prevents unconscious bias of including a stress-predominant MUI population in MUSA.

Behavioral therapy is first line management for urinary incontinence, either stress or urgency UI. In MUSA, women must report persistent symptoms despite at least one or more conservative treatments, (e.g. supervised behavioral/physical therapy).

After extensive discussion, the protocol committee ultimately decided that "failure" (inadequate response, intolerant of, or for whom they are contraindicated as determined by the physician) of medical therapy with anti-cholinergic and/or beta-mimetic medications would be an eligibility criterion per Botox label indications.

Based on the above rationale, the MUSA inclusion/exclusion criteria are as follows:

4.2 Inclusion Criteria

1. Reporting at least "moderate bother" from UUI item on UDI
   a. "Do you experience urine leakage associated with a feeling of urgency?"
2. Reporting at least "moderate bother" from SUI item on UDI
   a. "Do you experience urine leakage related to physical activity, coughing, or sneezing"
3. Diagnosis of SUI defined by a positive cough stress test (CST) or UDE within the past 18 months If participant does not demonstrate SUI during CMG they must demonstrate SUI through a cough stress test or other comparable valsalva maneuver to be eligible.
4. Presence of UUI on bladder diary with $\geq$ 4 Urgency IE/3-day diary
5. Urinary symptoms $\geq$3 months
6. Persistent symptoms despite at least one or more conservative treatments (e.g. supervised behavioral therapy, physical therapy) as determined adequate by the physician.
7. Inadequate response to oral overactive bladder medications (including anti-cholinergic and/or beta-mimetic medication) unless patient is
   a. intolerant of oral overactive bladder medications, or
   b. oral overactive bladder medications are contraindicated as determined by the treating provider.
8. Urodynamics within past 18 months prior to enrollment or done after enrollment, prior to randomization.

9. Demonstrates understanding (or have caregiver demonstrate understanding) to perform clean intermittent self-catheterization.
   Provider and patient review CISC process and patient (and/or caregiver) demonstrates understanding to the satisfaction of the provider.

## 4.3 Exclusion Criteria

1. Anterior or apical compartment prolapse at or beyond the hymen (≥0 on POPQ), regardless if patient is symptomatic
   a. Women with anterior or apical prolapse above the hymen (<0) who do not report vaginal bulge symptoms will be eligible
2. Planned concomitant surgery for anterior vaginal wall or apical prolapse ≥ 0
   b. Women undergoing only rectocele repair or other repair unrelated to anterior or apical compartment are eligible
3. Women undergoing hysterectomy for any indication will be excluded
4. Active pelvic organ malignancy
5. Age <21 years
6. Pregnant or plans for future pregnancy in next 6 months, or within 12 months post-partum
7. Post-void residual >150 cc on 2 occasions within the past 6 months, or current catheter use
8. Participation in other trial that may influence results of this study
9. Unevaluated hematuria
10. Prior sling, synthetic mesh for prolapse, implanted nerve stimulator for urinary incontinence
    Women with known Burch or Marshall-Marchetti-Krantz (MMK) are excluded.
11. Spinal cord injury or advanced/severe neurologic conditions including Multiple Sclerosis, Parkinsons, Myasthenia Gravis, Charcot-Marie-Tooth
12. Women on overactive bladder medication/therapy will be eligible after 3 week wash-out period
13. Non-ambulatory
14. History of serious adverse reaction to synthetic mesh
15. Not able to complete study assessments per clinician judgment, or not available for 6 month follow-up
16. Diagnosis of and/or history of bladder pain or chronic pelvic pain
17. Women who had intravesical Botox injection within the past 12 months
18. Women who have undergone anterior or apical pelvic organ prolapse repair within the past 6 months

## 4.4 Screening for Eligibility

It is anticipated that participants will be recruited from PFDN Clinical Site practices. Women with MUI will be offered the full range of treatment options consistent with routine practice including expectant management, pelvic floor muscle therapy, behavioral therapy, medication, and possibly surgery. Patients who are considered candidates for either Botox A injection or mid-urethral sling by their physician will be offered participation in MUSA.

Subjects will be identified as MUSA candidates by their physician. Subjects will be approached by study personnel consistent with local IRB requirements. Enrollment will occur after written and/or verbal consent. If the participant accepts participation in MUSA, the UDI will be administered to confirm at least moderate bother from both SUI and UUI and the coordinator will confirm documentation of SUI by either CST or UDE within the past 18 months, and UI symptoms for at least 3 months. She will be instructed on how to complete the voiding diary.

To address the issue of overactive bladder medication use, these subjects will be required to have a washout of 3 weeks prior to completing the voiding diary. The anticholinergic medication with the longest half-life currently on the market is Vesicare with a half-life of 45-68 hours. Therefore, by 1week there should be negligible amounts in the bloodstream and by 2 weeks the drug would be completely out of the system. Therefore, 3 weeks should be adequate time for washout and this time period is consistent with prior PFDN studies (ABC trial,[35] ESTEEM,[49] ROSETTA[33]). In addition, because we are highly interested in what happens to OAB outcomes, subjects will need to remain off overactive bladder medication for the 6 month duration of the study to allow accurate assessment of these symptoms following treatment (See statistical analysis plan for details). Subjects who take overactive bladder medication during the study period will be considered as having "additional treatment". Every effort will be made to schedule the patient's surgery within 2 months from enrollment (see Section 4.6, Appointment scheduling below).

## 4.5 Baseline Visit

At the baseline visit, the voiding diary will be reviewed to ensure that entries are clear and interpretable. If the first baseline voiding diary is not acceptable, the subject will be allowed one more attempt. If the second baseline voiding diary is not acceptable, the subject will not be eligible for the trial.

Once eligibility is confirmed, pre-treatment information will be obtained including:

- Demographics – age, race/ethnicity, education level
- Medical history – prior urinary incontinence procedures and treatments, prior pelvic surgeries, comorbidities, smoking, medications
- Physical examination – Body mass index, pelvic organ prolapse quantification (POPQ)
- Questionnaires – self-administered

## 4.6 Appointment Scheduling and Randomization

Once patients are enrolled and randomized, the intervention (either Botox A or mid-urethral sling) should be scheduled within 8 weeks from randomization. If a participant is randomized but does not undergo the intervention within 8 weeks, the patient can be retained in the study at the investigator's discretion until the planned intervention is scheduled, with the planned intervention date serving as Time 0 for calculating windows for follow up visits and phone calls. If surgery is rescheduled but does not occur, then the last

planned date of intervention will be Time 0. If the participant decides against the intervention but later changes her mind, the date of the intervention that did occur will be Time 0.

Postoperatively, all subjects will return and/or be contacted for visits at 2 weeks, 3 months, and 6 months. All subjects will return for a clinic visit at 2 weeks after Botox A or mid-urethral sling intervention to check PVR. PROs will be administered at baseline, 3 months, and 6 months post-intervention. (See Assessment Table 9).

# 5. DESCRIPTION OF STUDY INTERVENTIONS

## 5.1 Mid-urethral Sling Procedure

In MUSA, we plan to allow retropubic as well as transobturator full length slings. This is based on the findings of the TOMUS and Barber's equivalence trials in which these approaches demonstrated equivalence for improving objective success of SUI.[47,53] In these studies, both approaches also showed similar results for subjective success of SUI, persistent UUI or de novo UUI. A recent Cochrane review that evaluated data of 12,113 women undergoing mid-urethral sling procedures in 81 clinical trials has demonstrated the safety and efficacy of mid-urethral slings from several different manufacturers.[94]

"Mini-sling" and "single-incision" sling will not be allowed due to potential risk for higher failure rate.[95]

Key aspects of the procedure will be standardized across surgeons and sites.

### *5.1.1 Surgeon Certification*

To address the issue of surgeon certification for mid-urethral slings and to ensure standardized training of all surgeons, all "certified surgeons" will have performed a minimum of 20 mid-urethral slings of any type, including 5 of the mid-urethral sling allowed in MUSA that the surgeon will be using in the study. The site PI must sign off that each participating surgeon has met the criteria.

### *5.1.2 Standardization of Mid-urethral Sling Procedures*

Detailed standardization of the surgical procedure will be developed and will include the following key points:

1. The participating surgeon must be present and scrubbed for key portions of the procedure. Residents and fellows may participate in procedures as is standard for each Clinical Site
2. All subjects will receive preoperative intravenous antibiotic prophylaxis. The choice of antibiotic will be determined by each surgeon.
3. Deep vein thrombosis prophylaxis is required for all participants. The choice of prophylaxis will be determined by each surgeon.
4. Any concomitant native tissue procedures must be declared prior to randomization. Per exclusion criteria, women clinically requiring anterior vaginal prolapse or apical repairs are ineligible.
5. Tensioning of the sling will be performed in a fashion to ensure that it is a tension-free technique. This can include either by placing a blunt instrument between the sling and the urethra, or by folding a small knuckle of mesh in a Babcock clamp or similar method during tensioning.

### *5.1.3 Need for Postoperative Mid-urethral Sling Revision*

To address the issue of postoperative sling revision, the team developed a plan for several potential scenarios which may require the surgeon to revise the sling, detailed below. Women who undergo a sling revision will all be considered as having "additional treatment" regardless of indication. The reason for sling revision will be documented.

1. Urinary retention / incomplete bladder emptying (abnormal PVR) – An abnormal post-void residual is defined as PVR > 150 cc in this protocol (consistent with exclusion criteria). This is a known complication after mid-urethral sling. In the ESTEEM trial, the rate of sling revision was low (only one subject in each arm underwent sling revision). Similarly, Barber's trial which included 70% women with MUI, the sling revision rate was 0-1%, which is also consistent with the TOMUS trial. For retention/incomplete emptying, the postoperative management and need for sling revision will be left up to the surgeon's clinical judgment.
2. Worsening OAB/lower urinary tract symptoms with a **normal** PVR – it is possible that some women may experience worsening OAB symptoms immediately postoperatively. It is unclear from the literature in which women such symptoms may be transient and ultimately resolve once postoperative recovery is complete, or in which women it will persist and/or worsen over time. Therefore, for women with a normal PVR complaining of worsening OAB symptoms, sling revision will be deferred until 3 months postoperatively. This will provide important information about the natural course of these symptoms in the immediate postoperative period following sling. If after 3 months the patient desires sling revision due to worsening OAB symptoms, the surgeon can perform the procedure based on his/her clinical judgment. There is no evidence to support any potential harm by delaying sling revision in a woman with OAB symptoms and a normal PVR.
3. Persistent SUI symptoms – For women who have persistent SUI symptoms, sling revision/replacement can be performed after 3 months based on the surgeon's clinical judgment.

## 5.2 Botox Injection

### *5.2.1 Surgeon Certification*

To address the issue of surgeon certification for Botox intradetrusor injection and to ensure standardized training of all surgeons, all "certified surgeons" administering injections will be required to view a short instructional video demonstrating optimal techniques and detailing sites of drug injection in a standardized manner. In addition, "certified surgeons" at each site must have previously performed a total of 10 intradetrusor injection procedures.

### *5.2.2 Standardization of Botox Injection Procedures*

Detailed standardization of the surgical procedure will be developed and will include the following key points:

1. The Coordinator will ensure that the subject has been instructed regarding the proper technique for clean intermittent self-catheterization (CISC) and will confirm that the subject or designated caretaker is able to perform the task.

2. Within one week prior to injection, a confirming urine pregnancy test will be performed in all premenopausal women. A urine dip to rule out infection will also be performed. If the urine dip is positive the participant will be treated for a urinary tract infection and rescheduled within 4 weeks for Botox A® injection visit.

3. The bladder will be catheterized and 50 ml of 2% lidocaine placed in the bladder and 10 ml of 2% lidocaine jelly in the urethra. The subject will be asked to lie on their left and right lateral decubitus for several minutes at a time to ensure diffuse anesthetic effect on the bladder. Cystoscopic surveillance of the bladder will be used to confirm normality.

4. Cystoscopy will be performed with a 12 or 30-degree lens and rigid or flexible scope. A 22 gauge disposable needle, which is passed through the cystoscopic channel, will be used. Approximately 100-200 ml of total fluid will be instilled during cystoscopy to allow adequate visualization of the entire bladder urothelium. Botulinum toxin A will be prepared by dissolving 100 units of botulinum toxin A into 10 ml of injectable saline. Indigo carmine or methylene blue 0.1 ml will be added to each syringe of botulinum toxin A. The treating physician will inject a total of 10 ml of the Botox A® into approximately 15 to 20 different detrusor muscle sites under direct visualization. Injections will be spread out to equally cover the posterior bladder wall and dome, but spare the bladder trigone and ureteral orifices. The anterior bladder dome will be not be injected secondary to technical difficulties associated with injecting this area cystoscopically.

5. The Coordinator will record peri-procedural events and complications. The instilled fluid will be left in the bladder after the injections are complete. The subject will remain in a post-procedure area until a spontaneous void occurs. The patient must have a spontaneous void before going home. All subjects will receive a single dose of antibiotics orally per standard of care immediately after injection and subjects will take antibiotics orally once daily for 3 days post injection.

### *5.2.3 Criteria for Botox Reinjections*

Participants with persistent bothersome UUI at 3 months after their initial 100 unit Botox A injection may receive one additional injection of 100 units Botox A between 3 months and 6 months after initial injection. Eligibility for repeat Botox A injection are:

1. Persistent bothersome UUI as determined by reporting at least “moderate bother” from UUI item on UDI: “Do you experience urine leakage associated with a feeling of urgency, that is a strong sensation of needing to go to the bathroom?”
2. Continued UI bother based on the Patient Global Symptom Control (PGSC): “My current treatment is giving me adequate control of my urinary leakage”, score of ≤ 3
3. Participant desires additional treatment with Botox A
4. No UTI as determined by the health care provider on the day of the PGSC evaluation
5. PVR < 200 ml, per Botox A guidelines for UUI
6. No medical contraindication for the procedure as determined by the physician
7. Continuing to meet study inclusion/exclusion criteria (except inclusion items 2, 3, 4 7 and 8 and exclusion item 7).

A maximum of 1 additional 100 unit Botox A injection will be allowed between 3 and 6 months (total of 2 injections of 100 units Botox A). No injections will be given at an interval less than 3 months and the participant will have a post injection visit at 2 weeks (+/- 1 week) following injection.

## 5.3 Patient Management and Follow-Up

### *5.3.1 Baseline Procedures*

In addition to information collected to determine eligibility and standardized questionnaires, the following information will be obtained for all randomized patients by chart review or patient report:

a. Demographic information: age, race, ethnicity, insurance status, education

b. Medical history: vaginal parity, comorbidities, height, weight, prior pelvic surgeries, medications, estrogen status, previous treatments for pelvic floor disorders

c. Social history: tobacco use

d. Pelvic, rectal exam, and neurological examination will be done if indicated. POP-Q, post-void residual, urinary cough stress test will be documented.

e. Standardized urodynamic evaluation (UDE) will be performed before study interventions – there are no specific UDE parameters that determine eligibility for this trial beyond diagnosis of SUI defined by a positive CST or UDE. Small retrospective studies suggest that UDE parameters may be helpful in predicting outcomes after surgery in women with MUI and were predictors of failure in ESTEEM.[23,96] Therefore, patients in MUSA will undergo UDE testing, primarily to allow evaluation of variables that may predict clinical outcome. If patients have already undergone UDE prior to enrollment, results performed within 18 months of enrollment will be allowed if studies meet MUSA standards. For those women who have not, UDE will be performed at baseline.

f. Patient-reported outcomes and questionnaires – includes UDI, IIQ, EQ5D, PGI-I, PGI-S, OAB-q, OAB-SAT-q, PISQ-IR, SF-36, PGSC.

### *5.3.2 Postoperative Visits and Procedures*

Patients will undergo clinical and PRO assessments at 2-weeks, 3 months, and 6 months postoperatively (See **Table 9**). The primary outcome will be at 6 months.

## Table 9: Timeline of Visits, Events, and Data Collection

| | **Screening** | **Baseline** | **Procedure (Botox or Sling)** ‡, ** | **1-2 Week Visit** *‡ | **Telephone Call** | **3 Month Visit** | **Botox reinjection** ‡, ** | **1-2 Week Visit** *‡ | **Telephone Call** | **6 Month Visit†** |
|---|---|---|---|---|---|---|---|---|---|---|
| Window | 1-4 w | 1-4 w | | | 1w before 3m visit | 2 w pre – 4 w post | | | 1w before 6m visit | 2 w pre – 4 w post |
| Consent | X | | | | | | | | | |
| Coordinator visit | X | X | | X | X | X | | X | X | X |
| Hx/PE | X | X | | | | X | | | | X |
| Urodynamics** | | X | | | | | | | | |
| Medication and other treatment audit | X | X | | | | X | | | | X |
| Preg test~** | | | X | | | | X | | | |
| Urine dip~** | X | X | X | X | | X | X | X | | X |
| PVR | X | X | | X | | X | | X | | X |
| Adverse Events | | | X | X | | X | X | X | | X |
| PGSC | | | | | | X | | | | X |
| UDI | | X | | | | X | | | | X |
| Bladder diary | | X | | | Diary reminder | X | | | Diary reminder | X |
| PRO/Questionnaires Varies by milestone. | | X | | X | | X | | X | | X |

† primary outcome visit
* all subjects will undergo a post-op visit for PVR/urine dip 2 weeks after Botox A or sling procedure
‡ standard of care visit
** must be done in clinic

# 6. STATISTICAL CONSIDERATIONS

## 6.1 Sample size and Power

MUSA will randomize 146 women to treatment with Botox or Sling in a 1:1 ratio. The sample size was derived to provide sufficient power for the primary Aim (comparison of Botox versus sling after 6 months) assessed for the primary outcome (change in UDI total score).

### *6.1.1 Primary Aim*

**Assumptions**

For the primary outcome (UDI total score change from baseline at 6 months), alpha was set at 0.05 and power was set at 90%.

Sample size estimates were based on an evaluation of minimum important differences (MIDs) and corresponding standard deviation (SD) for the UDI total score from the ESTEEM study of MUI, with support from other published populations of women with urge and stress incontinence; however, the populations on which those MIDs were based might differ from the target population of mixed incontinence for MUSA. Thus, our goal was to power the study to detect a statistically significant difference between groups in change from baseline in UDI total score at 6 months for the UDI MID determined from ESTEEM as that will be a clinically important difference in our population. **Table 6** summarizes the published MIDs for ESTEEM, BE-DRI, and ATLAS. To use the most accurate standard deviation applicable to MUSA, we calculated the SD based on a mixed effects repeated measures model with adjustment for baseline UDI score and clinical site rather than the SD of the observed data in order to account for improvements in variance estimation from the covariate adjustment planned for MUSA. To most closely match MUSA, this analysis model included all 3-month and 6-month ESTEEM data points after any retreatments and excludes patients with <4 Urge IEs/3 days. The MID for ESTEEM UDI-Total is 26, and the SD=46.5 from the analysis model, resulting in an effect size = 0.56.

Total number of randomized patients required to meet planned power estimates were adjusted for the assumed patient discontinuation rate at 6 months. ROSETTA and ESTEEM discontinuation rates are displayed in **Table 10**. Based on average estimates from ROSETTA and ESTEEM, we plan for 5% discontinuation rate by 6 months.

**Table 10: 6 Month Discontinuation Rates from ROSETTA and ESTEEM**

| Study | Population | 6-Month Discontinuation Rate |
|---|---|---|
| ROSETTA | UUI | 10/386=2.6% |
| ESTEEM sling arm | MUI | 16/235 = 6.8% |
| Plan for MUSA | MUI | 5% |

**Sample Size and Power Calculation**

Sample size estimates were generated using nQuery Advisor under the assumptions outlined above. The evaluable sample size required at 6 months for 90% power for the primary total UDI outcome is 69 per arm. Because data from ROSETTA and ESTEEM suggests that follow-up at 6 months should be about 95%, a total sample size of 73 randomized per treatment arm is selected (69/0.95 =73), for a total sample size of 146 randomized participants. These calculations are provided in **Table 11**.

**Table 11: MUSA Power Calculation for N=146 (73 Per Treatment Arm)**

| UDI component | Primary or Secondary Outcome | MID | SD [a] | Effect Size | Alpha | Power at 6 months | Randomized Sample Size per Group [b] |
|---|---|---|---|---|---|---|---|
| UDI-Total | Primary | 26 (ESTEEM) | 46.5 | 0.56 | 0.05 | 90% | 73 |
| UDI-Total | Primary | 26 (ESTEEM) | 46.5 | 0.56 | 0.05 | 85% | 62 |
| UDI-Total | Primary | 26 (ESTEEM) | 46.5 | 0.56 | 0.05 | 80% | 55 |
| | | | | | | | |
| UDI-Total | Primary | 23 | 46.5 | 0.50 | 0.05 | 82% | 73 |

[a] SD: standard deviation is estimated from ESTEEM participants with at least 4 Urge IEs/3 days using a repeated measures linear model for change from baseline to 3 and 6 months with adjustment for baseline value and clinical site.
[b] Assumes 5% of randomized participants discontinue by 6 months estimated by average of ROSETTA and ESTEEM

**Sensitivity to MID Assumptions for the Primary Outcome**

With the sample size of 146 subjects and the planned standard deviation, our study will have 82% power to detect a statistically significant difference in UDI-total score at $p<0.05$ if the true difference is as small as 23 points (0.50 effect size) between groups at 6 months (**Table 11**). Thus, the planned sample size may allow for analyses to assess whether the true MID in this population is smaller than 26.

## 6.2 Statistical Analysis Plan

### *6.2.1 Analysis Populations and Management of Study Discontinuations Prior to Receiving Treatment*

It is possible that some women in both groups may cancel their surgical procedure due to personal or other reasons. Consistent with analyses for prior PFDN protocols (SUPeR, ASPIRe, ROSETTA), non-treated patients will be removed from evaluation of efficacy in a modified intent-to-treat population.

### *6.2.2 Primary Aim*

**Primary Outcome**

The mean change from baseline in UDI total score will be compared between groups at 6 months at the $p<0.05$ statistical significance level using a mixed effects analysis of covariance model for repeated measures (MMRM)with adjustment for baseline UDI score, and randomization stratification factors site and age group. The model will include fixed effect categorical factors for treatment group, visit (3 and 6 months), and treatment × visit interaction with a heterogeneous Toeplitz covariance pattern for the within-subject repeated measures. Estimates, p-values and 95% confidence intervals will be presented for treatment group comparisons at 6 moths (primary) and 3 months (secondary).

This model assumes missing data due to a missed visit or early study discontinuation is missing at random. If more than 10% of participants are missing their 6-month score, then a sensitivity analysis will be performed using multiple imputation under a variety of missing data patterns to be specified in the statistical analysis plan in order to assess the robustness of the primary results to missing data. .

We will report whether change in total UDI score from baseline to 6 months is statistically significantly different between groups. If the difference is statistically significant, the potential clinical significance of the difference will be discussed. We recognize that our sample size may allow us to find a difference between groups that is statistically significant yet smaller than published MID for total UDI score for women with MUI. However, published MIDs were calculated based on populations that may be somewhat different from the one targeted for enrollment in MUSA, and an exploratory aim of MUSA is to explore whether the MID in this population differs from previously published values.

**Secondary Outcomes**

There are two secondary outcomes for the primary aim. The mean change from baseline in UDI-stress and UDI-irritative subscores will be compared between groups at 6 months. Each of these outcomes will be evaluated using an alpha level of 0.05, with no adjustment for multiple comparisons. The statistical analysis of the two secondary outcomes will be identical to the analysis described above for the UDI total score.

### *6.2.3 Exploratory Aims*

All statistical evaluations of the exploratory aims are for descriptive purposes and p-values resulting from these analyses will not be evaluated for statistical significance.

1. **Secondary urinary outcomes:** To compare treatment with Botox A to treatment with mid-urethral sling for improving the number of urinary incontinence episodes on bladder diary after 6 months.

   Changes from baseline in bladder diary outcomes will be calculated and analyzed using the methods described for the analysis of the primary outcome. For urinary frequency, women reporting on average >8 voids/24 hours at baseline will be considered symptomatic, and normalization of voiding frequency will be defined as $\leq$ 8 voids/24 hours at 6 months. A 50% improvement will be defined as a reduction by half in the number of voids that patients had at baseline. The number of women who had normalization of voiding frequency and 50% improvement at 6 months will be compared between groups separately and collectively. We will also assess the proportion of women who had worsening of urinary frequency (includes women who developed de novo frequency and those who worsened). These dichotomous outcomes will be analyzed using logistic regression, controlling for the design effects of stratification by site and age group.

2. **Predictors of poor treatment response:** To develop models to identify baseline predictors of change of MUI, OAB, and SUI outcomes measured using the UDI between baseline and 6 months post-treatment

Regression models will be created to identify predictors of change from baseline to 6 months for UDI total score and stress and irritative subscale scores. Potential predictors will include age, diary parameters such as number of UUI episodes/3 days, functional bladder capacity, stress and irritative bother severity at baseline, type of urinary incontinence (stress-predominant or urgency-predominant: see definition below), and co-existing anterior vaginal wall prolapse. The relationship between potential predictors and outcomes will be

explored in models that include one predictor plus stratification factors (site and age group). Predictive models will be constructed using backward selection of predictors. The impact of collinearity between predictors will be assessed and the final model modified as necessary.

The definition of the type of urinary incontinence, stress- or urgency predominant will be based on 2 questions of the Urogenital Distress Inventory,

1. "Do you experience urine leakage associated with a feeling of urgency?"
2. "Do you experience urine leakage related to physical activity, coughing, or sneezing?"

Based on our inclusion criteria, patients must report at least "moderate" bother on both questions.
If the subject reports greater bother on the stress question, she will be classified as stress-predominant MUI. If the subject reports greater bother on the urgency question, she will be classified as urgency-predominant MUI. If the subject reports equal bother on both questions, she will be classified as "balanced" MUI.

3. **Quality of life and global impression:** To compare quality of life outcomes and Patient Global Impression-Improvement (PGI-I), Patient Global Impression-Severity (PGI-S) between groups randomized to Botox versus sling after 6 months.

   Changes from baseline in all quality of life PRO instruments and global impression will be calculated and analyzed using the methods described for the analysis of the primary outcome

4. **Safety and additional treatments:** To describe rates of reoperation (sling revision) after mid-urethral sling and intermittent catheterization due to voiding dysfunction/partial urinary retention after Botox injection, to compare the proportion of women in each group with UTI and recurrent UTI, rates of other serious and non-serious adverse events, and to compare the proportion of women in each group initiating additional (off protocol) treatment other than Botox and mid-urethral sling for SUI and/or OAB, and the types of additional treatment (behavioral therapy, medications, other).

   Groups will be summarized with n and percent, and rates will be compared with chi-square tests.

5. **Cost-effectiveness analysis:** To determine the cost effectiveness of Botox A injection versus mid-urethral sling for the treatment of MUI symptoms at 6 months.

   Differential mean costs and differential mean QALYs between the two treatment groups will be estimated using multiple regression analysis. Specifically, a generalized linear model with appropriate link function (e.g., log-link) and response probability distribution (e.g., gamma distribution) will be used to analyze costs due to the potential skewness and heteroscedasticity of medical expenditure data, while an ordinary least squares regression will be used for analyzing QALY data. The models will account for stratification factors of study site and age group, as well as other characteristics of the subjects that are found to differ significantly between the groups. When estimating QALYs, we will also adjust for subjects' baseline utility scores to account for potential imbalance in baseline utility between the two treatment groups.[97]

   We will calculate the incremental cost-effectiveness ratio (ICER), which is the differential mean costs divided by the differential mean QALYs between the two groups, to assess the additional costs associated with each additional QALY gained. Our base case analysis will be conducted based on subjects with complete data. Sensitivity analysis will be conducted to include subjects with incomplete data using the multiple imputation method. Non-parametric bootstrapping

resampling technique will be used to derive the 95% confidence interval for the ICER.[98] In addition, cost-effectiveness acceptability curve (CEAC) will be generated to illustrate the likelihood that one treatment is more cost-effective than the other with various ceiling cost-effectiveness ratios.

We plan to conduct supplemental analyses using alternative outcome measures, such as incremental cost per reduction in UIE/day.

The cost-effectiveness evaluations will be conducted as within-trial comparisons. A decision analytic model will also be developed from trial data to evaluate the trajectory of the cost-effectiveness ratio.

6. **UDI MID:** To explore MIDs for UDI total score and stress and irritative subscores for this MUI population

   MIDs will be calculated using anchor- and distribution-based approaches. Potential anchors include the PGSC, global impression of change, and incontinence episodes from the bladder diary.

### 6.3 Interim Data Monitoring

Safety outcomes will be assessed at each DSMB meeting. This will include the need for sling revision due to worsening OAB symptoms.

Since we expect to enroll MUSA within 7 months, and since the primary outcome is attained at 6 months following surgery, no formal interim analyses of efficacy outcomes will be performed. At each meeting, the DSMB will be presented with information about enrollment, participant adverse events, and outcome data attainment (for example, the percent of expected clinic visits that have been completed) to allow them to determine that the study is making reasonable progress.

## 7. ETHICAL CONCERNS/SAFETY

### 7.1 Ethical Concerns

As discussed in the background section, current clinical practice varies with respect to treatment of MUI and likely reflects training and experiential bias. Although treatment with behavioral modifications and Kegel exercises have been described as effective first line treatments for mild stress, urgency, and mixed urinary incontinence, many patients go on to request further therapy for their condition. For moderate symptoms of SUI or UUI, additional therapeutic options are generally offered based on treatment paradigms geared toward each of these conditions. When patients have MUI, clinicians must decide which component (the SUI or the UUI) should be addressed first. There is very little evidence to support a defined treatment strategy in this patient population and most recommendations are based on expert opinion. The only way to test the superiority of one approach over another is in the setting of a randomized clinical trial. We have carefully designed this trial to balance the risks and benefits to subjects. All patients in this study will have already been offered more conservative therapies. We will be assured that women will have previously tried behavioral and/or pelvic floor therapy because it is an inclusion criterion. The potential benefits of the Botox A intervention are improvement of UUI and MUI symptoms while the risks are minimal. The benefit of mid-urethral sling for patients with SUI is well documented and several studies, including ESTEEM, have also documented an improvement in OAB and MUI symptoms following sling.

## 7.2 Informed Consent

Subjects will be clinically examined as part of screening and to ensure eligibility for the study. Those subjects who are candidates for mid-urethral and willing to undergo sling surgery and Botox A treatment for MUI will be approached for enrollment into the trial. Clinical and research staff will describe the study in detail and answer any questions the subject may have. Documented informed consent for trial participation will be obtained at that time. A common template for the research informed consent form will be used by all of the clinical sites, modifying the content or format as necessary to meet the requirements of their respective institutional human subjects committees. This study must be approved by the IRBs at the clinical sites and DCC before study implementation.

## 7.3 Data Safety Monitoring Board

The National Institutes of Health has set up a Data Safety Monitoring Board (DSMB) to oversee all PFDN studies, including this study. Members of the DSMB are independent of the study investigators and represent Urology, Urogynecology and Biostatistics, as well as having a lay member. The DSMB meets 3 times per year, or more frequently if requested by the Chair, either in person or by teleconference. This protocol has been approved by the DSMB prior to implementation. Safety and efficacy data will be assessed in a descriptive manner at each DSMB meeting without formal statistical tests.

## 7.4 Reporting of Serious Adverse Events

Each clinical investigator is responsible for reporting serious adverse events (SAEs) to the IRB per their IRB guidelines at their institution, and to the DCC. The DCC Safety Specialist reviews and summarizes the SAE per DCC SAE reporting procedures for the PFDN. The DSMB reviews all SAEs.

## 7.5 Side Effects/Safety

### *7.5.1 Mid-urethral Sling*

Sling surgery is a commonly performed operation for the treatment of SUI and MUI. Like all surgical interventions it has the risk of bleeding, infection, and injury to surrounding structures. In addition, the sling procedure utilizes polypropylene mesh which can introduce additional risk of mesh complication. These include vaginal mesh extrusion, mesh infection, and bladder or urethral mesh erosion. Complications specific to sling placement include bladder perforation, retropubic hematoma, obturator nerve or vessel injury, groin pain, worsening incontinence, and worsening OAB. The FDA has recently issued guidelines on the use of surgical mesh and has recommended it only be used by trained surgeons. All surgeons participating in this study will be specifically trained to use surgical mesh.

### *7.5.2 Botox A*

Botox A detrusor injection is a commonly performed procedure for the treatment of UUI and MUI. Like all surgical interventions it has the risk of bleeding, infection, and injury to surrounding structures.

A recent comprehensive review identified 44 original research studies that reported on 16 different conditions treated with multiple treatments with Botox A ®. The conditions treated with Botox A ® were varied and included such disorders such as achalasia, blepharospasm, cervical dystonia, cerebral palsy, esophageal spasm, hemifacial spasm, laryngeal dystonia, oromandibular dystonia, strabismus in addition to

detrusor overactivity. Botox A® has been FDA approved since 1989 for only some of the conditions it is currently been used for. According to the American Society of Plastic Surgery, 7.4 million cosmetic botulinum toxin A injections were administered in 2018. This figure was up from 787,000 in 2000. http://www.plasticsurgery.org.

### *7.5.3 Botox A Contraindications*[102]

BOTOX is contraindicated in patients who are hypersensitive to any botulinum toxin product or to any of the components in the formulation. In the presence of infection at the proposed injection site(s). For intradetrusor injection in patients with a urinary tract infection; or in patients with urinary retention or post-void residual (PVR) urine volume >200 mL who are not routinely performing clean intermittent self-catheterization (CIC).

### *7.5.4 Botox A Warnings and Precautions*[102]

**Lack of Interchangeability Between Botulinum Toxin Products**

The potency Units of BOTOX are specific to the preparation and assay method utilized. They are not interchangeable with other preparations of botulinum toxin products and, therefore, units of biological activity of BOTOX cannot be compared to nor converted into units of any other botulinum toxin products assessed with any other specific assay method [see Description].

**Spread of Toxin Effect**

Post-marketing safety data from BOTOX and other approved botulinum toxins suggest that botulinum toxin effects may, in some cases, be observed beyond the site of local injection. The symptoms are consistent with the mechanism of action of botulinum toxin and may include asthenia, generalized muscle weakness, diplopia, ptosis, dysphagia, dysphonia, dysarthria, urinary incontinence, and breathing difficulties. These symptoms have been reported hours to weeks after injection. Swallowing and breathing difficulties can be life threatening and there have been reports of death related to spread of toxin effects. The risk of symptoms is probably greatest in children treated for spasticity but symptoms can also occur in adults treated for spasticity and other conditions, and particularly in those patients who have an underlying condition that would predispose them to these symptoms. In unapproved uses and in approved indications, symptoms consistent with spread of toxin effect have been reported at doses comparable to or lower than doses used to treat cervical dystonia and spasticity. Patients or caregivers should be advised to seek immediate medical care if swallowing, speech or respiratory disorders occur.

No definitive serious adverse event reports of distant spread of toxin effect associated with BOTOX for blepharospasm at the recommended dose (30 Units and below), severe primary axillary hyperhidrosis at the recommended dose (100 Units), strabismus, or for chronic migraine at the labeled doses have been reported.

**Serious Adverse Reactions with Unapproved Use**

Serious adverse reactions, including excessive weakness, dysphagia, and aspiration pneumonia, with some adverse reactions associated with fatal outcomes, have been reported in patients who received BOTOX injections for unapproved uses. In these cases, the adverse reactions were not necessarily related to distant spread of toxin, but may have resulted from the administration of BOTOX to the site of injection and/or adjacent structures. In several of the cases, patients had pre-existing dysphagia or other significant disabilities. There is insufficient information to identify factors associated with an increased risk for adverse

reactions associated with the unapproved uses of BOTOX. The safety and effectiveness of BOTOX for unapproved uses have not been established.

**Hypersensitivity Reactions**

Serious and/or immediate hypersensitivity reactions have been reported. These reactions include anaphylaxis, serum sickness, urticaria, soft tissue edema, and dyspnea. If such a reaction occurs, further injection of BOTOX should be discontinued and appropriate medical therapy immediately instituted. One fatal case of anaphylaxis has been reported in which lidocaine was used as the diluent, and consequently the causal agent cannot be reliably determined.

**Increased Risk of Clinically Significant Effects with Pre-Existing Neuromuscular Disorders**

Individuals with peripheral motor neuropathic diseases, amyotrophic lateral sclerosis or neuromuscular junction disorders (e.g., myasthenia gravis or Lambert-Eaton syndrome) should be monitored when given botulinum toxin. Patients with known or unrecognized neuromuscular disorders or neuromuscular junction disorders may be at increased risk of clinically significant effects including generalized muscle weakness, diplopia, ptosis, dysphonia, dysarthria, severe dysphagia and respiratory compromise from therapeutic doses of BOTOX [see Warnings and Precautions].

**Dysphagia and Breathing Difficulties**

Treatment with BOTOX and other botulinum toxin products can result in swallowing or breathing difficulties. Patients with pre-existing swallowing or breathing difficulties may be more susceptible to these complications. In most cases, this is a consequence of weakening of muscles in the area of injection that are involved in breathing or oropharyngeal muscles that control swallowing or breathing [see Warnings and Precautions].

Deaths as a complication of severe dysphagia have been reported after treatment with botulinum toxin. Dysphagia may persist for several months, and require use of a feeding tube to maintain adequate nutrition and hydration. Aspiration may result from severe dysphagia and is a particular risk when treating patients in whom swallowing or respiratory function is already compromised.

Treatment with botulinum toxins may weaken neck muscles that serve as accessory muscles of ventilation. This may result in a critical loss of breathing capacity in patients with respiratory disorders who may have become dependent upon these accessory muscles. There have been post-marketing reports of serious breathing difficulties, including respiratory failure.

Patients with smaller neck muscle mass and patients who require bilateral injections into the sternocleidomastoid muscle for the treatment of cervical dystonia have been reported to be at greater risk for dysphagia. Limiting the dose injected into the sternocleidomastoid muscle may reduce the occurrence of dysphagia. Injections into the levator scapulae may be associated with an increased risk of upper respiratory infection and dysphagia.

Patients treated with botulinum toxin may require immediate medical attention should they develop problems with swallowing, speech or respiratory disorders. These reactions can occur within hours to weeks after injection with botulinum toxin [see Warnings and Precautions].

**Pulmonary Effects of Botox in Patients with Compromised Respiratory Status Treated for Detrusor Overactivity Associated with a Neurologic Condition**

In a double-blind, placebo-controlled, parallel group study in adult patients with detrusor overactivity associated with a neurologic condition and restrictive lung disease of neuromuscular etiology [defined as FVC 50-80% of predicted value in patients with spinal cord injury between C5 and C8, or MS] the event rate in change of Forced Vital Capacity ≥15% or ≥20% was generally greater in patients treated with BOTOX than in patients treated with placebo (see **Table 12**).

**Table 12: Number and Percent of Patients Experiencing at Least a 15% or 20% Decrease in FVC From Baseline at Week 2, 6, 12 Post-injection with BOTOX or Placebo**

| BOTOX | 200 Units | | Placebo | |
|---|---|---|---|---|
| | >15% | >20% | >15% | >20% |
| Week 2 | 0/15 (0%) | 0/15 (0%) | 1/11 (9%) | 0/11 (0%) |
| Week 6 | 2/13 (15%) | 1/13 (8%) | 0/12 (0%) | 0/12 (0%) |
| Week 12 | 0/12(0%) | 0/12 (0%) | 0/7 (0%) | 0/7 (0%) |

**Autonomic Dysreflexia in Patients Treated for Detrusor Overactivity Associated with a Neurologic Condition**

Autonomic dysreflexia associated with intradetrusor injections of BOTOX could occur in patients treated for detrusor overactivity associated with a neurologic condition and may require prompt medical therapy. In clinical trials, the incidence of autonomic dysreflexia was greater in patients treated with BOTOX 200 Units compared with placebo (1.5% versus 0.4%, respectively).

**Urinary Tract Infections in Patients with Overactive Bladder**

BOTOX increases the incidence of urinary tract infection [see Adverse Reactions]. Clinical trials for overactive bladder excluded patients with more than 2 UTIs in the past 6 months and those taking antibiotics chronically due to recurrent UTIs. Use of BOTOX for the treatment of overactive bladder in such patients and in patients with multiple recurrent UTIs during treatment should only be considered when the benefit is likely to outweigh the potential risk.

**Urinary Retention in Patients Treated for Bladder Dysfunction**

Due to the risk of urinary retention, treat only patients who are willing and able to initiate catheterization post-treatment, if required, for urinary retention.

In patients who are not catheterizing, post-void residual (PVR) urine volume should be assessed within 2 weeks post-treatment and periodically as medically appropriate up to 12 weeks, particularly in patients with multiple sclerosis or diabetes mellitus. Depending on patient symptoms, institute catheterization if PVR urine volume exceeds 200 mL and continue until PVR falls below 200 mL. Instruct patients to contact their physician if they experience difficulty in voiding as catheterization may be required.

The incidence and duration of urinary retention is described below for patients with overactive bladder and detrusor overactivity associated with a neurologic condition who received BOTOX or placebo injections.

## Overactive Bladder

In double-blind, placebo-controlled trials in patients with OAB, the proportion of subjects who initiated clean intermittent catheterization (CIC) for urinary retention following treatment with BOTOX or placebo is shown in **Table 13**. The duration of post-injection catheterization for those who developed urinary retention is also shown.

**Table 13: Proportion of Patients Catheterizing for Urinary Retention and Duration of Catheterization Following an Injection in Double-blind, Placebo-Controlled Clinical Trials in OAB**

| Timepoint | BOTOX 100 Units (N=552) | Placebo (N=542) |
|---|---|---|
| **Proportion of Patients Catheterizing for Urinary Retention** | | |
| At any time during complete treatment cycle | 6.5% (n=36) | 0.4% (n=2) |
| **Duration of Catheterization for Urinary Retention (Days)** | | |
| Median | 63 | 11 |
| Min, Max | 1, 214 | 3, 18 |

Patients with diabetes mellitus treated with BOTOX were more likely to develop urinary retention than those without diabetes, as shown in **Table 14**.

**Table 14: Proportion of Patients Experiencing Urinary Retention Following an Injection in Double-blind, Placebo-controlled Clinical Trials in OAB According to History of Diabetes Mellitus**

| | Patients with Diabetes | | Patients without Diabetes | |
|---|---|---|---|---|
| | BOTOX 100 Units (N=81) | Placebo (N=69) | BOTOX 100 Units (N=526) | Placebo (N=516) |
| Urinary retention | 12.3% (n=10) | 0 | 6.3% (n=33) | 0.6% (n=3) |

## Detrusor Overactivity Associated with a Neurologic Condition

In two double-blind, placebo-controlled trials in patients with detrusor overactivity associated with a neurologic condition (NDO-1 and NDO-2), the proportion of subjects who were not using clean intermittent catheterization (CIC) prior to injection and who subsequently required catheterization for urinary retention following treatment with BOTOX 200 Units or placebo is shown in **Table 15**. The duration of post-injection catheterization for those who developed urinary retention is also shown.

**Table 15: Proportion of Patients Not Using CIC at Baseline and then Catheterizing for Urinary Retention and Duration of Catheterization Following an Injection in Double-blind, Placebo-controlled Clinical Trials**

| Timepoint | BOTOX 200 Units (N=108) | Placebo (N=104) |
|---|---|---|
| **Proportion of Patients Catheterizing for Urinary Retention** | | |
| At any time during complete treatment cycle | 30.6% (n=33) | 6.7% (n=7) |
| **Duration of Catheterization for Urinary Retention (Days)** | | |
| Median | 289 | 358 |
| Min, Max | 1, 530 | 2, 379 |

## Human Albumin and Transmission of Viral Diseases

This product contains albumin, a derivative of human blood. Based on effective donor screening and product manufacturing processes, it carries an extremely remote risk for transmission of viral diseases and variant Creutzfeldt-Jakob disease (vCJD). There is a theoretical risk for transmission of Creutzfeldt-Jakob disease (CJD), but if that risk actually exists, the risk of transmission would also be considered extremely remote. No cases of transmission of viral diseases, CJD or vCJD have ever been identified for licensed albumin or albumin contained in other licensed products.

### *7.5.5 Adverse Reactions to Botox A*[102]

The following adverse reactions to BOTOX (onabotulinumtoxinA) for injection are discussed in greater detail in other sections of the labeling:

- Spread of Toxin Effects
- Serious Adverse Reactions with Unapproved Use
- Hypersensitivity Reactions
- Increased Risk of Clinically Significant Effects with Pre-Existing Neuromuscular Disorders
- Dysphagia and Breathing Difficulties
- Pulmonary Effects of BOTOX in Patients with Compromised Respiratory Status Treated for Spasticity or for Detrusor Overactivity associated with a Neurologic Condition
- Corneal Exposure and Ulceration in Patients Treated with BOTOX for Blepharospasm
- Retrobulbar Hemorrhages in Patients Treated with BOTOX for Strabismus
- Bronchitis and Upper Respiratory Tract Infections in Patients Treated for Spasticity
- Autonomic Dysreflexia in Patients Treated for Detrusor Overactivity associated with a Neurologic Condition
- Urinary Tract Infections in Patients with Overactive Bladder
- Urinary Retention in Patients Treated for Bladder Dysfunction

## Botox A Clinical Trials Experience

Because clinical trials are conducted under widely varying conditions, the adverse reaction rates observed in the clinical trials of a drug cannot be directly compared to rates in the clinical trials of another drug and may not reflect the rates observed in clinical practice.

BOTOX and BOTOX Cosmetic contain the same active ingredient in the same formulation, but with different labeled Indications and Usage. Therefore, adverse reactions observed with the use of BOTOX Cosmetic also have the potential to be observed with the use of BOTOX.

In general, adverse reactions occur within the first week following injection of BOTOX and, while generally transient, may have a duration of several months or longer. Localized pain, infection, inflammation, tenderness, swelling, erythema, and/or bleeding/bruising may be associated with the injection. Symptoms associated with flu-like symptoms (e.g., nausea, fever, myalgia) have been reported after treatment. Needle-related pain and/or anxiety may result in vasovagal responses (including syncope, hypotension), which may require appropriate medical therapy.

Local weakness of the injected muscle(s) represents the expected pharmacological action of botulinum toxin. However, weakness of nearby muscles may also occur due to spread of toxin

***Overactive Bladder***
**Table 16** presents the most frequently reported adverse reactions in double-blind, placebo-controlled clinical trials for overactive bladder occurring within 12 weeks of the first BOTOX treatment.

**Table 16: Adverse Reactions Reported by ≥2% of BOTOX treated Patients and More Often than in Placebo-treated Patients Within the First 12 Weeks after Intradetrusor Injection, in Double-blind, Placebo-controlled Clinical Trials in Patients with OAB**

| Adverse Reactions | BOTOX 100 Units (N=552) % | Placebo (N=542) % |
|---|---|---|
| Urinary tract infection | 18 | 6 |
| Dysuria | 9 | 7 |
| Urinary retention | 6 | 0 |
| Bacteriuria | 4 | 2 |
| Residual urine volume* | 3 | 0 |

*Elevated PVR not requiring catheterization. Catheterization was required for PVR >350 mL regardless of symptoms, and for PVR >200 mL to <350 mL with symptoms (e.g., voiding difficulty).

A higher incidence of urinary tract infection was observed in patients with diabetes mellitus treated with BOTOX 100 Units and placebo than in patients without diabetes, as shown in **Table 17**.

**Table 17: Proportion of Patients Experiencing Urinary Tract Infection following an Injection in Double-blind, Placebo-controlled Clinical Trials in OAB according to history of Diabetes Mellitus**

| | Patients with Diabetes | | Patients without Diabetes | |
|---|---|---|---|---|
| | BOTOX 100 Units (N=81) % | Placebo (N=69) % | BOTOX 100 Units (N=526) % | Placebo (N=516) % |
| Urinary tract infection (UTI) | 31 | 12 | 26 | 10 |

The incidence of UTI increased in patients who experienced a maximum post-void residual (PVR) urine volume >200 mL following BOTOX injection compared to those with a maximum PVR <200 mL following BOTOX injection, 44% versus 23%, respectively. No change was observed in the overall safety profile with repeat dosing during an open-label, uncontrolled extension trial.

### *7.5.6 Immunogenicity of Botox A*[102]

As with all therapeutic proteins, there is a potential for immunogenicity. The detection of antibody formation is highly dependent on the sensitivity and specificity of the assay. Additionally, the observed incidence of antibody (including neutralizing antibody) positivity in an assay may be influenced by several factors including assay methodology, sample handling, timing of sample collection, concomitant medications, and underlying disease. For these reasons, comparison of the incidence of antibodies to onabotulinumtoxinA in the studies described below with the incidence of antibodies in other studies or to other products may be misleading.

In a long term, open-label study evaluating 326 cervical dystonia patients treated for an average of 9 treatment sessions with the current formulation of BOTOX, 4 (1.2%) patients had positive antibody tests. All 4 of these patients responded to BOTOX therapy at the time of the positive antibody test. However, 3 of these patients developed clinical resistance after subsequent treatment, while the fourth patient continued to respond to BOTOX therapy for the remainder of the study.

One patient among the 445 hyperhidrosis patients (0.2%), two patients among the 380 adult upper limb spasticity patients (0.5%), and no patients among 406 migraine patients with analyzed specimens developed the presence of neutralizing antibodies.

In overactive bladder patients with analyzed specimens from the two phase 3 studies and the open-label extension study, neutralizing antibodies developed in 0 of 954 patients (0.0%) while receiving BOTOX 100 Unit doses and 3 of 260 patients (1.2%) after subsequently receiving at least one 150 Unit dose. Response to subsequent BOTOX treatment was not different following seroconversion in these three patients.

In detrusor overactivity associated with neurologic condition patients with analyzed specimens in the drug development program (including the open-label extension study), neutralizing antibodies developed in 3 of 300 patients (1.0%) after receiving only BOTOX 200 Unit doses and 5 of 258 patients (1.9%) after receiving at least one 300 Unit dose. Following development of neutralizing antibodies in these 8 patients, 4 continued to experience clinical benefit, 2 did not experience clinical benefit, and the effect on the response to BOTOX in the remaining 2 patients is not known.

The data reflect the patients whose test results were considered positive for neutralizing activity to BOTOX in a mouse protection assay or negative based on a screening ELISA assay or mouse protection assay.

Formation of neutralizing antibodies to botulinum toxin type A may reduce the effectiveness of BOTOX treatment by inactivating the biological activity of the toxin. The critical factors for neutralizing antibody formation have not been well characterized. The results from some studies suggest that BOTOX injections at more frequent intervals or at higher doses may lead to greater incidence of antibody formation. The potential for antibody

### *7.5.7 Post-Marketing Experience of Botox A*[102]

The following adverse reactions have been identified during post-approval use of BOTOX. Because these reactions are reported voluntarily from a population of uncertain size, it is not always possible to reliably estimate their frequency or establish a causal relationship to drug exposure. These reactions include: abdominal pain; alopecia, including madarosis; anorexia; brachial plexopathy; denervation/muscle atrophy; diarrhea; dry eye; hyperhidrosis; hypoacusis; hypoaesthesia; localized muscle twitching; malaise; paresthesia; peripheral neuropathy; radiculopathy; erythema multiforme, dermatitis psoriasiform, and psoriasiform eruption; strabismus; tinnitus; and visual disturbances.

There have been spontaneous reports of death, sometimes associated with dysphagia, pneumonia, and/or other significant debility or anaphylaxis, after treatment with botulinum toxin.

There have also been reports of adverse events involving the cardiovascular system, including arrhythmia and myocardial infarction, some with fatal outcomes. Some of these patients had risk factors including cardiovascular disease. The exact relationship of these events to the botulinum toxin injection has not been established.

New onset or recurrent seizures have also been reported, typically in patients who are predisposed to experiencing these events. The exact relationship of these events to the botulinum toxin injection has not been established.

## 7.6 Drug Interactions with Botox A[102]

### *7.6.1 Aminoglycosides and Other Agents Interfering with Neuromuscular Transmission*

Co-administration of BOTOX and aminoglycosides or other agents interfering with neuromuscular transmission (e.g., curare-like compounds) should only be performed with caution as the effect of the toxin may be potentiated.

### *7.6.2 Anticholinergic Drugs*

Use of anticholinergic drugs after administration of BOTOX may potentiate systemic anticholinergic effects.

### *7.6.3 Other Botulinum Neurotoxin Products*

The effect of administering different botulinum neurotoxin products at the same time or within several months of each other is unknown. Excessive neuromuscular weakness may be exacerbated by administration of another botulinum toxin prior to the resolution of the effects of a previously administered botulinum toxin.

### *7.6.4 Muscle Relaxants*

Excessive weakness may also be exaggerated by administration of a muscle relaxant before or after administration of BOTOX.

## 7.7 Use of Botox A in Specific Populations[102]

### *7.7.1 Pregnancy*

**Risk Summary**
There are no studies or adequate data from postmarketing surveillance on the developmental risk associated with use of BOTOX in pregnant women. In animal studies, administration of BOTOX during pregnancy resulted in adverse effects on fetal growth (decreased fetal weight and skeletal ossification) at clinically relevant doses, which were associated with maternal toxicity. In the U.S. general population, the estimated background risk of major birth defects and miscarriages in clinically recognized pregnancies is 2-4% and 15-20%, respectively. The background risk of major birth defects and miscarriage for the indicated populations is unknown.

**Animal Data**
When BOTOX (4, 8, or 16 Units/kg) was administered intramuscularly to pregnant mice or rats two times during the period of organogenesis (on gestation days 5 and 13), reductions in fetal body weight and decreased fetal skeletal ossification were observed at the two highest doses. The no-effect dose for developmental toxicity in these studies (4 Units/kg) is approximately equal to the human dose of 400 Units, on a body weight basis (Units/kg).

When BOTOX was administered intramuscularly to pregnant rats (0.125, 0.25, 0.5, 1, 4, or 8 Units/kg) or rabbits (0.063, 0.125, 0.25, or 0.5 Units/kg) daily during the period of organogenesis (total of 12 doses in rats, 13 doses in rabbits), reduced fetal body weights and decreased fetal skeletal ossification were observed at the two highest doses in rats and at the highest dose in rabbits. These doses were also associated with significant maternal toxicity, including abortions, early deliveries, and maternal death. The developmental no-effect doses in these studies of 1 Unit/kg in rats and 0.25 Units/kg in rabbits are less than the human dose of 400 Units, based on Units/kg.

When pregnant rats received single intramuscular injections (1, 4, or 16 Units/kg) at three different periods of development (prior to implantation, implantation, or organogenesis), no adverse effects on fetal development were observed. The developmental no-effect level for a single maternal dose in rats (16 Units/kg) is approximately 2 times the human dose of 400 Units, based on Units/kg.

### *7.7.2 Lactation*

**Risk Summary**
There are no data on the presence of BOTOX in human or animal milk, the effects on the breastfed infant, or the effects on milk production. The developmental and health benefits of breastfeeding should be considered along with the mother's clinical need for BOTOX and any potential adverse effects on the breastfed infant from BOTOX or from the underlying maternal conditions.

### *7.7.3 Geriatric Use*

Of the 2145 adult patients in placebo-controlled clinical studies of BOTOX for the treatment of spasticity, 33.5% were 65 or older, and 7.7% were 75 years of age or older. No overall differences in safety were observed between elderly patients and adult patients younger than 65 years of age.

In clinical studies of BOTOX across other indications, no overall differences in safety were observed between elderly patients and younger adult patients, with the exception of Overactive Bladder (see below). Other reported clinical experience has not identified differences in responses between the elderly and younger adult patients, but greater sensitivity of some older individuals cannot be ruled out.

### *7.7.4 Overactive Bladder*

Of 1242 overactive bladder patients in placebo-controlled clinical studies of BOTOX, 41.4% were 65 years of age or older, and 14.7% were 75 years of age or older. Adverse reactions of UTI and urinary retention were more common in patients 65 years of age or older in both placebo and BOTOX groups compared to younger patients (see **Table 18**). Otherwise, there were no overall differences in the safety profile following BOTOX treatment between patients aged 65 years and older compared to adult patients younger than 65 years of age in these studies.

**Table 18: Incidence of Urinary Tract Infection and Urinary Retention according to Age Group**

| | **<65 Years** | | **65 to 74 Years** | | **≥75 Years** | |
|---|---|---|---|---|---|---|
| **Adverse Reactions** | **BOTOX 100 Units (N=344) %** | **Placebo (N=348) %** | **BOTOX 100 Units (N=169) %** | **Placebo (N=151) %** | **BOTOX 100 Units (N=94) %** | **Placebo (N=86) %** |
| Urinary tract infection | 21 | 7 | 30 | 13 | 38 | 19 |

Observed effectiveness was comparable between these age groups in placebo-controlled clinical studies.

## 7.8 Overdosage of Botox A[102]

Excessive doses of BOTOX (onabotulinumtoxinA) for injection may be expected to produce neuromuscular weakness with a variety of symptoms.

Symptoms of overdose are likely not to be present immediately following injection. Should accidental injection or oral ingestion occur or overdose be suspected, the person should be medically supervised for several weeks for signs and symptoms of systemic muscular weakness which could be local, or distant from the site of injection [see Boxed Warning and Warnings and Precautions (5.2, 5.6)]. These patients should be considered for further medical evaluation and appropriate medical therapy immediately instituted, which may include hospitalization.

If the musculature of the oropharynx and esophagus are affected, aspiration may occur which may lead to development of aspiration pneumonia. If the respiratory muscles become paralyzed or sufficiently weakened, intubation and assisted respiration may be necessary until recovery takes place. Supportive care could involve the need for a tracheostomy and/or prolonged mechanical ventilation, in addition to other general supportive care.

In the event of overdose, antitoxin raised against botulinum toxin is available from the Centers for Disease Control and Prevention (CDC) in Atlanta, GA. However, the antitoxin will not reverse any botulinum toxin-induced effects already apparent by the time of antitoxin administration. In the event of suspected or actual

cases of botulinum toxin poisoning, please contact your local or state Health Department to process a request for antitoxin through the CDC. If you do not receive a response within 30 minutes, please contact the CDC directly at 1-770-488-7100. More information can be obtained at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5232a8.htm.

### 7.9 Adverse Event Reporting

Adverse events are defined as untoward medical events that are temporally related to participation in a clinical study, regardless of whether they are causally related to the study. Adverse events will be collected during the course of this study and reported to the DSMB as described above.

Adverse events will be reported in a manner consistent with the requirements outlined in the NICHD Clinical Research Policy Guidance Document. That document specifies which adverse events be reported to the Office for Human Research Protections (OHRP), and to the FDA, if FDA-regulated products such as a device, drug, or biologic are used. Consistent with that policy, adverse events will be reported in a manner consistent with OHRP and FDA regulations.

## 8. FEASIBILITY

The proposed MUI study population has already failed conservative management and are interested in more effective therapies. We have taken care to have highly effective treatments for both UUI (Botox A) and SUI (mid-urethral sling) available if needed in both arms, in a clinical efficacy trial design with inclusion criteria that are not overly strict. In addition, a population similar to this population has been previously recruited within the PFDN for ESTEEM; therefore, we do not anticipate particular difficulty in recruitment of MUI patients as encountered in MIMOSA.[60]

## 9. REFERENCES

1. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program in Women (CPW) Research Group. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 1994;3:291-306.
2. Melville JL, Katon W, Delaney K, Newton K. Urinary incontinence in US women: a population-based study. Archives of internal medicine 2005;165:537-42.
3. Karram MM, Bhatia NN. Management of coexistent stress and urge urinary incontinence. Obstetrics and gynecology 1989;73:4-7.
4. Stewart WF, Van Rooyen JB, Cundiff GW, et al. Prevalence and burden of overactive bladder in the United States. World journal of urology 2003;20:327-36.
5. Agarwal A, Eryuzlu LN, Cartwright R, et al. What is the most bothersome lower urinary tract symptom? Individual- and population-level perspectives for both men and women. European urology 2014;65:1211-7.
6. Monz B, Chartier-Kastler E, Hampel C, et al. Patient characteristics associated with quality of life in European women seeking treatment for urinary incontinence: results from PURE. European urology 2007;51:1073-81; discussion 81-2.
7. Dooley Y, Lowenstein L, Kenton K, FitzGerald M, Brubaker L. Mixed incontinence is more bothersome than pure incontinence subtypes. International urogynecology journal and pelvic floor dysfunction 2008;19:1359-62.
8. Coyne KS, Zhou Z, Thompson C, Versi E. The impact on health-related quality of life of stress, urge and mixed urinary incontinence. BJU international 2003;92:731-5.

9. Minassian VA, Devore E, Hagan K, Grodstein F. Severity of urinary incontinence and effect on quality of life in women by incontinence type. Obstetrics and gynecology 2013;121:1083-90.
10. Barber MD, Spino C, Janz NK, et al. 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 2009;200:580.e1-7.
11. Haylen BT, de Ridder D, Freeman RM, et al. An International Urogynecological Association (IUGA)/International Continence Society (ICS) joint report on the terminology for female pelvic floor dysfunction. Neurourology and urodynamics 2010;29:4-20.
12. Brubaker L, Stoddard A, Richter H, et al. Mixed incontinence: comparing definitions in women having stress incontinence surgery. Neurourology and urodynamics 2009;28:268-73.
13. Herzog AR, Diokno AC, Brown MB, Normolle DP, Brock BM. Two-year incidence, remission, and change patterns of urinary incontinence in noninstitutionalized older adults. Journal of gerontology 1990;45:M67-74.
14. Brubaker L, Lukacz ES, Burgio K, et al. Mixed incontinence: comparing definitions in non-surgical patients. Neurourology and urodynamics 2011;30:47-51.
15. Petros PE. Mixed urinary incontinence--time to uncouple urgency from stress? International urogynecology journal 2011;22:919-21.
16. Langer R, Lipshitz Y, Halperin R, Pansky M, Bukovsky I, Sherman D. Long-Term (10-15 years) follow-up after Burch colposuspension for urinary stress incontinence. International urogynecology journal and pelvic floor dysfunction 2001;12:323-6; discussion 6-7.
17. Jung SY, Fraser MO, Ozawa H, et al. Urethral afferent nerve activity affects the micturition reflex; implication for the relationship between stress incontinence and detrusor instability. The Journal of urology 1999;162:204-12.
18. Koonings P, Bergman A, Ballard CA. Combined detrusor instability and stress urinary incontinence: where is the primary pathology? Gynecologic and obstetric investigation 1988;26:250-6.
19. Dekhtiar YM, Kostyev FI, Zacheslavsky OM, Kuznietsov DO. Urodynamic characteristics of lower urinary tract of patients with idiopathic overactive bladder. Urology annals 2019;11:83-6.
20. Dietz HP, Clarke B. The urethral pressure profile and ultrasound imaging of the lower urinary tract. International urogynecology journal and pelvic floor dysfunction 2001;12:38-41.
21. Sinha S, Lakhani D, Singh VP. Cough associated detrusor overactivity in women with urinary incontinence. Neurourology and urodynamics 2019;38:920-6.
22. Kulseng-Hanssen S, Moe K, Schiotz HA. How often does detrusor overactivity cause urinary leakage during a stress test in women with mixed urinary incontinence? International urogynecology journal 2013;24:1537-41.
23. Paick JS, Ku JH, Kim SW, Oh SJ, Son H, Shin JW. Tension-free vaginal tape procedure for the treatment of mixed urinary incontinence: significance of maximal urethral closure pressure. The Journal of urology 2004;172:1001-5.
24. Richter HE, Moalli P, Amundsen CL, et al. Urinary Biomarkers in Women with Refractory Urgency Urinary Incontinence Randomized to Sacral Neuromodulation versus OnabotulinumtoxinA Compared to Controls. The Journal of urology 2017;197:1487-95.
25. Chai TC, Richter HE, Moalli P, et al. Inflammatory and tissue remodeling urinary biomarkers before and after mid urethral sling surgery for stress urinary incontinence. The Journal of urology 2014;191:703-9.
26. Dumoulin C, Cacciari LP, Hay-Smith EJC. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women. The Cochrane database of systematic reviews 2018;10:Cd005654.

27. Balk EM, Rofeberg VN, Adam GP, Kimmel HJ, Trikalinos TA, Jeppson PC. Pharmacologic and Nonpharmacologic Treatments for Urinary Incontinence in Women: A Systematic Review and Network Meta-analysis of Clinical Outcomes. Annals of internal medicine 2019.
28. Burgio KL, Kraus SR, Menefee S, et al. Behavioral therapy to enable women with urge incontinence to discontinue drug treatment: a randomized trial. Annals of internal medicine 2008;149:161-9.
29. Sexton CC, Notte SM, Maroulis C, et al. Persistence and adherence in the treatment of overactive bladder syndrome with anticholinergic therapy: a systematic review of the literature. International journal of clinical practice 2011;65:567-85.
30. Gray SL, Anderson ML, Dublin S, et al. Cumulative use of strong anticholinergics and incident dementia: a prospective cohort study. JAMA internal medicine 2015;175:401-7.
31. Kelleher C, Hakimi Z, Zur R, et al. Efficacy and Tolerability of Mirabegron Compared with Antimuscarinic Monotherapy or Combination Therapies for Overactive Bladder: A Systematic Review and Network Meta-analysis. European urology 2018;74:324-33.
32. Drake MJ, Nitti VW, Ginsberg DA, et al. Comparative assessment of the efficacy of onabotulinumtoxinA and oral therapies (anticholinergics and mirabegron) for overactive bladder: a systematic review and network meta-analysis. BJU international 2017;120:611-22.
33. Amundsen CL, Richter HE, Menefee SA, et al. OnabotulinumtoxinA vs Sacral Neuromodulation on Refractory Urgency Urinary Incontinence in Women: A Randomized Clinical Trial. Jama 2016;316:1366-74.
34. Amundsen CL, Komesu YM, Chermansky C, et al. Two-Year Outcomes of Sacral Neuromodulation Versus OnabotulinumtoxinA for Refractory Urgency Urinary Incontinence: A Randomized Trial. European urology 2018;74:66-73.
35. Visco AG, Brubaker L, Richter HE, et al. Anticholinergic therapy vs. onabotulinumtoxina for urgency urinary incontinence. The New England journal of medicine 2012;367:1803-13.
36. Chapple C, Sievert KD, MacDiarmid S, et al. OnabotulinumtoxinA 100 U significantly improves all idiopathic overactive bladder symptoms and quality of life in patients with overactive bladder and urinary incontinence: a randomised, double-blind, placebo-controlled trial. European urology 2013;64:249-56.
37. Nitti VW, Ginsberg D, Sievert KD, et al. Durable Efficacy and Safety of Long-Term OnabotulinumtoxinA Treatment in Patients with Overactive Bladder Syndrome: Final Results of a 3.5-Year Study. The Journal of urology 2016;196:791-800.
38. Dmochowski R, Chapple C, Nitti VW, et al. Efficacy and safety of onabotulinumtoxinA for idiopathic overactive bladder: a double-blind, placebo controlled, randomized, dose ranging trial. The Journal of urology 2010;184:2416-22.
39. Nitti VW, Dmochowski R, Herschorn S, et al. OnabotulinumtoxinA for the treatment of patients with overactive bladder and urinary incontinence: results of a phase 3, randomized, placebo controlled trial. The Journal of urology 2013;189:2186-93.
40. Tincello DG, Kenyon S, Abrams KR, et al. Botulinum toxin a versus placebo for refractory detrusor overactivity in women: a randomised blinded placebo-controlled trial of 240 women (the RELAX study). European urology 2012;62:507-14.
41. Sahai A, Khan MS, Dasgupta P. Efficacy of botulinum toxin-A for treating idiopathic detrusor overactivity: results from a single center, randomized, double-blind, placebo controlled trial. The Journal of urology 2007;177:2231-6.
42. Brubaker L, Richter HE, Visco A, et al. Refractory idiopathic urge urinary incontinence and botulinum A injection. The Journal of urology 2008;180:217-22.

43. Flynn MK, Amundsen CL, Perevich M, Liu F, Webster GD. Outcome of a randomized, double-blind, placebo controlled trial of botulinum A toxin for refractory overactive bladder. The Journal of urology 2009;181:2608-15.
44. Kobashi KC, Albo ME, Dmochowski RR, et al. Surgical Treatment of Female Stress Urinary Incontinence: AUA/SUFU Guideline. The Journal of urology 2017;198:875-83.
45. Abrams P, Andersson KE, Birder L, et al. Fourth International Consultation on Incontinence Recommendations of the International Scientific Committee: Evaluation and treatment of urinary incontinence, pelvic organ prolapse, and fecal incontinence. Neurourology and urodynamics 2010;29:213-40.
46. Welk B, Baverstock RJ. The management of mixed urinary incontinence in women. Canadian Urological Association journal = Journal de l'Association des urologues du Canada 2017;11:S121-s4.
47. Richter HE, Albo ME, Zyczynski HM, et al. Retropubic versus transobturator midurethral slings for stress incontinence. The New England journal of medicine 2010;362:2066-76.
48. Jain P, Jirschele K, Botros SM, Latthe PM. Effectiveness of midurethral slings in mixed urinary incontinence: a systematic review and meta-analysis. International urogynecology journal 2011;22:923-32.
49. Sung VW, Borello-France D, Dunivan G, et al. Methods for a multicenter randomized trial for mixed urinary incontinence: rationale and patient-centeredness of the ESTEEM trial. International urogynecology journal 2016;27:1479-90.
50. Tahseen S, Reid P. Effect of transobturator tape on overactive bladder symptoms and urge urinary incontinence in women with mixed urinary incontinence. Obstetrics and gynecology 2009;113:617-23.
51. Barber MD, Kleeman S, Karram MM, et al. Risk factors associated with failure 1 year after retropubic or transobturator midurethral slings. American journal of obstetrics and gynecology 2008;199:666.e1-7.
52. Houwert RM, Venema PL, Aquarius AE, Bruinse HW, Roovers JP, Vervest HA. Risk factors for failure of retropubic and transobturator midurethral slings. American journal of obstetrics and gynecology 2009;201:202.e1-8.
53. Barber MD, Kleeman S, Karram MM, et al. Transobturator tape compared with tension-free vaginal tape for the treatment of stress urinary incontinence: a randomized controlled trial. Obstetrics and gynecology 2008;111:611-21.
54. 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 2005;193:103-13.
55. Richter HE, Litman HJ, Lukacz ES, et al. Demographic and clinical predictors of treatment failure one year after midurethral sling surgery. Obstetrics and gynecology 2011;117:913-21.
56. Nager CW, Sirls L, Litman HJ, et al. Baseline urodynamic predictors of treatment failure 1 year after mid urethral sling surgery. The Journal of urology 2011;186:597-603..
57. Shumaker SA, Wyman JF, Uebersax JS, McClish D, Fantl JA (1994) Health-related quality of life measures for women with urinary incontinence: the incontinence impact questionnaire and the urogenital distress inventory. Continence program in women (CPW) research group. Qual Life Res 3:291–306
58. Abdel-fattah M, Ramsay I, Pringle S, Hardwick C, Ali H. Evaluation of transobturator tapes (E-TOT) study: randomised prospective single-blinded study comparing inside-out vs. outside-in transobturator tapes in management of urodynamic stress incontinence: short term outcomes. European journal of obstetrics, gynecology, and reproductive biology 2010;149:106-11.

59. Abdel-Fattah M, Ramsay I, Pringle S, et al. Evaluation of transobturator tension-free vaginal tapes in management of women with recurrent stress urinary incontinence. Urology 2011;77:1070-5.
60. Brubaker L, Moalli P, Richter HE, et al. Challenges in designing a pragmatic clinical trial: the mixed incontinence -- medical or surgical approach (MIMOSA) trial experience. Clinical trials (London, England) 2009;6:355-64.
61. Miotla P, Futyma K, Cartwright R, et al. Effectiveness of botulinum toxin injection in the treatment of de novo OAB symptoms following midurethral sling surgery. International urogynecology journal 2016;27:393-8.
62. Denys P, Le Normand L, Ghout I, et al. Efficacy and safety of low doses of onabotulinumtoxinA for the treatment of refractory idiopathic overactive bladder: a multicentre, double-blind, randomised, placebo-controlled dose-ranging study. European urology 2012;61:520-9.
64. Richter HE, Amundsen CL, Erickson SW, et al. Characteristics Associated with Treatment Response and Satisfaction in Women Undergoing OnabotulinumtoxinA and Sacral Neuromodulation for Refractory Urgency Urinary Incontinence. The Journal of urology 2017;198:890-6.
65. Komesu YM, Amundsen CL, Richter HE, et al. Refractory urgency urinary incontinence treatment in women: impact of age on outcomes and complications. American journal of obstetrics and gynecology 2018;218:111.e1-.e9.
66. Trabuco EC, Klingele CJ, Weaver AL, McGree ME, Lightner DJ, Gebhart JB. Preoperative and postoperative predictors of satisfaction after surgical treatment of stress urinary incontinence. American journal of obstetrics and gynecology 2011;204:444.e1-6.
67. Maldonado PA, Kogutt BK, Wai CY. Patient satisfaction following midurethral sling surgeries. Current opinion in obstetrics & gynecology 2014;26:404-8.
68. Aydin S, Arioglu Aydin C, Ersan F. Prediction of Mid-Urethral Sling Failure with Clinical Findings and Urodynamics. Lower urinary tract symptoms 2017;9:89-93.
69. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. (Accessed June 16, 2009, at https://www.fda.gov/media/77832/download.)
70. Nager CW, Brubaker L, Litman HJ, et al. A randomized trial of urodynamic testing before stress-incontinence surgery. The New England journal of medicine 2012;366:1987-97.
71. Dyer KY, Xu Y, Brubaker L, et al. Minimum important difference for validated instruments in women with urge incontinence. Neurourology and urodynamics 2011;30:1319-24.
72. Goode PS, Burgio KL, Kraus SR, Kenton K, Litman HJ, Richter HE. Correlates and predictors of patient satisfaction with drug therapy and combined drug therapy and behavioral training for urgency urinary incontinence in women. International urogynecology journal 2011;22:327-34.
73. Lowenstein L, Kenton K, FitzGerald MP, Brubaker L. Clinically useful measures in women with mixed urinary incontinence. American journal of obstetrics and gynecology 2008;198:664.e1-3; discussion .e3-4.
74. Gleason JL, Parden AM, Jauk V, Ballard A, Sung V, Richter HE. Outcomes of midurethral sling procedures in women with mixed urinary incontinence. International urogynecology journal 2015;26:715-20.
75. Margolis MK, Fox KM, Cerulli A, Ariely R, Kahler KH, Coyne KS. Psychometric validation of the overactive bladder satisfaction with treatment questionnaire (OAB-SAT-q). Neurourology and urodynamics 2009;28:416-22.
76. Coyne KS, Matza LS, Thompson CL. The responsiveness of the Overactive Bladder Questionnaire (OAB-q). Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2005;14:849-55.

77. Avery KN, Bosch JL, Gotoh M, et al. Questionnaires to assess urinary and anal incontinence: review and recommendations. The Journal of urology 2007;177:39-49.
78. Coyne KS, Matza LS, Thompson CL, Kopp ZS, Khullar V. Determining the importance of change in the overactive bladder questionnaire. The Journal of urology 2006;176:627-32; discussion 32.
79. Coyne KS, Matza LS, Thompson C, Jumadilova Z, Bavendam T. The responsiveness of the OAB-q among OAB patient subgroups. Neurourology and urodynamics 2007;26:196-203.
80. Rogers RG, Coates KW, Kammerer-Doak D, Khalsa S, Qualls C. A short form of the Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-12). International urogynecology journal and pelvic floor dysfunction 2003;14:164-8; discussion 8.
81. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care 1992;30:473-83.
82. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of health economics 2002;21:271-92.
83. Yalcin I, Bump RC. Validation of two global impression questionnaires for incontinence. American journal of obstetrics and gynecology 2003;189:98-101.
84. Brooks R. EuroQol: the current state of play. Health policy (Amsterdam, Netherlands) 1996;37:53-72.
85. Harvie HS, Lee DD, Andy UU, Shea JA, Arya LA. Validity of utility measures for women with pelvic organ prolapse. American journal of obstetrics and gynecology 2018;218:119.e1-.e8.
86. Harvie HS, Shea JA, Andy UU, Propert K, Schwartz JS, Arya LA. Validity of utility measures for women with urge, stress, and mixed urinary incontinence. American journal of obstetrics and gynecology 2014;210:85.e1-6.
87. Harvie HS, Honeycutt AA, Neuwahl SJ, et al. Responsiveness and minimally important difference of SF-6D and EQ-5D utility scores for the treatment of pelvic organ prolapse. American journal of obstetrics and gynecology 2019;220:265.e1-.e11.
88. Longworth L, Bryan S. An empirical comparison of EQ-5D and SF-6D in liver transplant patients. Health economics 2003;12:1061-7.
89. Albo ME, Richter HE, Brubaker L, et al. Burch colposuspension versus fascial sling to reduce urinary stress incontinence. The New England journal of medicine 2007;356:2143-55.
90. Nager CW, Brubaker L, Daneshgari F, et al. Design of the Value of Urodynamic Evaluation (ValUE) trial: A non-inferiority randomized trial of preoperative urodynamic investigations. Contemporary clinical trials 2009;30:531-9.
91. Bosch JL, Cardozo L, Hashim H, Hilton P, Oelke M, Robinson D. Constructing trials to show whether urodynamic studies are necessary in lower urinary tract dysfunction. Neurourology and urodynamics 2011;30:735-40.
92. Nager CW, Kraus SR, Kenton K, et al. Urodynamics, the supine empty bladder stress test, and incontinence severity. Neurourology and urodynamics 2010;29:1306-11.
93. Rovner ES, Goudelocke CM. Urodynamics in the evaluation of overactive bladder. Current urology reports 2010;11:343-7.
94. Ford AA, Rogerson L, Cody JD, Aluko P, Ogah JA. Mid-urethral sling operations for stress urinary incontinence in women. The Cochrane database of systematic reviews 2017;7:Cd006375.
95. Nambiar A, Cody JD, Jeffery ST, Aluko P. Single-incision sling operations for urinary incontinence in women. The Cochrane database of systematic reviews 2017;7:Cd008709.
96. Paick JS, Ku JH, Shin JW, Son H, Oh SJ, Kim SW. Tension-free vaginal tape procedure for urinary incontinence with low Valsalva leak point pressure. The Journal of urology 2004;172:1370-3.
97. McCulloch CE, Searle SR. Generalized, linear, and mixed models. New York: John Wiley & Sons; 2001.
98. Efron B, Tibshirani R. An introduction to the bootstrap. New York: Chapman & Hall; 1994.

102 OnabotulinumtoxinA [package insert]. Madison, NJ: Allergan, Inc. (2019).

## 10. ADDENDUM - 12 Month Extension

**Amendment: Extension of the Treatment for Mixed Urinary Incontinence: Mid-urethral Sling vs. Botox A (MUSA) study for follow-up to 12 months**

**Background**

The primary purpose of the MUSA trial was to estimate the effect of intradetrusor injections of Botulinum toxin A (Botox A ®) compared to mid-urethral sling for the treatment of mixed urinary incontinence (MUI) symptoms in 146 women at 6 months. MUSA is a 1:1 randomized 2-arm clinical trial: intradetrusor injection of 100 units of Botox A, with a second dose after 3 months if needed, versus mid-urethral sling. **The primary aim** was to compare the effectiveness of intradetrusor injection of 100 units of Botox A to mid-urethral sling for change in MUI symptoms 6 months following treatment. **The primary Outcome** was change in severity of MUI symptoms 6 months post treatment measured using the Urogenital Distress Inventory (UDI) in patients randomized to either Botox A or mid-urethral sling. **The secondary outcomes** were 1) change in severity of stress urinary incontinence (SUI) and urgency urinary incontinence (UUI) symptoms 6 months post treatment measured using the stress and irritative subscales of the UDI and 2) change in MUI symptoms 3 months post treatment measured using the UDI total score.

Exploratory Aims included:

1. **Secondary urinary outcomes:** To compare treatment with Botox A to treatment with mid-urethral sling for improving the number of urinary incontinence episodes on bladder diary 6 months post-treatment.
2. **Predictors of poor treatment response:** To develop models to identify baseline predictors of change of MUI, OAB, and SUI outcomes measured using the UDI, between baseline and 6 months post-treatment.
3. **Quality of life and global impression:** To compare quality of life outcomes and Patient Global Impression-Improvement (PGI-I), Patient Global Impression-Severity (PGI-S) between groups randomized to Botox A versus mid-urethral sling 6 months post-treatment.
4. **Safety and additional treatments:** To describe rates of reoperation (sling revision) after mid-urethral sling and intermittent catheterization due to voiding dysfunction/partial urinary retention after Botox A detrusor injection, to compare the proportion of women in each group with UTI and recurrent UTI, rates of other serious and non-serious adverse events, and to compare the proportion of women in each group initiating additional (off protocol) treatment other than Botox A and mid-urethral sling for SUI and/or OAB.
5. **Cost-effectiveness analysis:** To determine the cost effectiveness of Botox A injection versus mid-urethral sling for the treatment of MUI symptoms on an intent-to-treat basis 6 months post-treatment.
6. **UDI MID:** To explore MIDs for UDI total score and stress and irritative subscores for this MUI population.

The MUSA study was initially designed as a 12 month study, however due to constraints related to funding and completing the study within the current PFDN cycle, the study was revised to a 6 month outcome. The DSMB recommended a 12 month follow up. 12 month outcomes are common in midurethral sling surgical

trials and a similar outcome to the ESTEEM study. The PFND bridge year funding now allows us the opportunity to extend the MUSA trial to 12 month outcomes.

MUSA recruitment began in July 2020 and was estimated to be completed by May 2021, by the end of the current PFDN cycle. The bridge year would extend until June 2022 and allow completion of the 12 month MUSA outcome within the bridge year. The first MUSA subject 9-month and 12-month windows would open Apr 2021and August 2021, respectively. The current cycle ends in June 2021, if the MUSA 12 month follow-up needs to be discontinued due to time or financial reasons, there would be a potential of 9, 9-month calls and no 12-month calls completed before June 2021.

**Proposed Extension of Study**

This amendment proposes to continue to follow all the MUSA subjects from 6 to 12 months. **The 6-12 month extension would be observational.** Participants would be allowed to pursue other clinical treatments after 6 months, including the alternative arm treatment. **The primary MUSA study outcome remains at 6 months. The 12-month aims, outcomes, outcome measures, and analysis are similar to the existing MUSA study at 6 months.**

**12-month aim:** To compare the effectiveness of intradetrusor injection of 100 units of Botox A to mid-urethral sling for change in MUI symptoms 12 months following treatment.

**Primary Outcome:** Change in severity of MUI symptoms 12 months post treatment measured using the UDI

**Secondary Outcomes:** Change in severity of SUI and UUI symptoms 12 months post treatment measured using the stress and irritative subscales of the UDI

Additional study visits for the 12-month extension include a 9-month call and a 12-month call. The first MUSA subject will reach the 9-month call window in April 2021. The table below represents the outcome measures that will be collected at each MUSA call or visit. The existing MUSA outcome measures and data collection forms remain the same for the 9-month call (excluding the bladder diary, PVR, and urine dip) and 12-month call (excluding the PVR, and urine dip).

**Table 1: MUSA Visit Schedule and Data Collection Forms**

| **MUSA Study Visit** | **3 M Visit** | **6 M Visit** | **9 M Call** | **12 M Call** |
|---|---|---|---|---|
| **MUSA Data Collection Forms** | | | | |
| UDI | **X** | **X** | **X** | **X** |
| Bladder Diary | **X** | **X** | | **X** |
| PVR – Clinic Visit | **X** | **X** | | |
| Urine Dip | **X** | **X** | | |
| Medical History – follow up | **X** | **X** | **X** | **X** |
| Overactive Bladder Questionnaire-Symptom subscale (OAB-q) | **X** | **X** | **X** | **X** |

| | | | | |
|---|---|---|---|---|
| Incontinence Impact Questionnaire (IIQ) | X | X | X | X |
| Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ) | X | X | X | X |
| European Quality of Life-5 Dimensions (EQ-5D) | X | X | X | X |
| Short Form 36 (SF-36)/ Short Form 6D (SG-6D) | X | X | X | X |
| Patient Global Impression of Severity (PGI-S). | X | X | X | X |
| Lost Productivity / Costs | X | X | X | X |
| Overactive Bladder treatment satisfaction (OAB-SAT-q) | X | X | X | X |
| Patient Global Impression of Improvement (PGI-I) | X | X | X | X |
| PGSC | X | X | X | X |

**Sample Size Estimate**

For the MUSA study, sample size estimates were based on an evaluation of minimum important differences (MIDs) and corresponding standard deviation (SD) for the UDI total score from the ESTEEM study, with support from other published populations of women with urge and stress incontinence. For the primary outcome (UDI total score change from baseline at 6 months), alpha was set at 0.05 and power was set at 90%. We assumed a discontinuation rate of 5% at 6 months based on ROSETTA and ESTEEM, for a total sample size of 146 (Table 2). Assuming an additional 5% discontinuation rate between 6 and 12 month, we would still have 88% power for the 12 month UDI total score outcome.

**Table 2: MUSA Power Calculation for N=146 (73 Per Treatment Arm)**

| UDI component | Primary or Secondary Outcome | MID | SD [a] | Effect Size | Alpha | Power at 6 months | Randomized Sample Size per Group [b] |
|---|---|---|---|---|---|---|---|
| UDI-Total | Primary | 26 (ESTEEM) | 46.5 | 0.56 | 0.05 | 90% | 73 |
| UDI-Total | Primary | 26 (ESTEEM) | 46.5 | 0.56 | 0.05 | 85% | 62 |
| UDI-Total | Primary | 26 (ESTEEM) | 46.5 | 0.56 | 0.05 | 80% | 55 |

[a] SD: standard deviation is estimated from ESTEEM participants with at least 4 Urge IEs/3 days using a repeated measures linear model for change from baseline to 3 and 6 months with adjustment for baseline value and clinical site.
[b] Assumes 5% of randomized participants discontinue by 6 months estimated by average of ROSETTA and ESTEEM

**Statistical Analysis and Feasibility**

The analysis at 12 months will be similar to planned 6 month analysis. The primary MUSA study outcome remains at 6 months.

The discontinuation rate of ESTEEM and ROSETTA were 12% and 9%, respectively at 12 months. An additional participant incentive of $75 will be given at the 12 month call. Given that MUSA participants are allowed to pursue other clinical treatments after 6 months, including the alternative arm treatment, retention

from 6 to 12 months should not exceed 10%. The 9 and 12 month calls will be halted if we determine that it is not feasible to collect sufficient data due to time or budget constraints

Statistical Analysis Plan

***Primary outcome:*** The mean change from baseline in UDI total score will be compared between groups at 12 months at the $p<0.05$ statistical significance level using a mixed effects analysis of covariance model for repeated measures (MMRM)with adjustment for baseline UDI score, and randomization stratification factors site and age group. The model will include fixed effect categorical factors for treatment group, visit and treatment × visit interaction with a heterogeneous Toeplitz covariance pattern for the within-subject repeated measures. Estimates, p-values and 95% confidence intervals will be presented for treatment group comparisons at 12 months.

This model assumes missing data due to a missed visit or early study discontinuation is missing at random. If more than 10% of participants are missing their 12-month score, then a sensitivity analysis will be performed using multiple imputation under a variety of missing data patterns to be specified in the statistical analysis plan in order to assess the robustness of the primary results to missing data.

We will report whether change in total UDI score from baseline to 12 months is statistically significantly different between groups. If the difference is statistically significant, the potential clinical significance of the difference will be discussed. We recognize that our sample size may allow us to find a difference between groups that is statistically significant yet smaller than published MID for total UDI score for women with MUI. However, published MIDs were calculated based on populations that may be somewhat different from the one targeted for enrollment in MUSA, and an exploratory aim of MUSA is to explore whether the MID in this population differs from previously published values.

***Secondary outcomes*:** There are two secondary outcomes for the primary aim. The mean change from baseline in UDI-stress and UDI-irritative subscores will be compared between groups at 12 months. Each of these outcomes will be evaluated using an alpha level of 0.05, with no adjustment for multiple comparisons. The statistical analysis of the two secondary outcomes will be identical to the analysis described above for the UDI total score.

***Exploratory aims:*** All statistical evaluations of the exploratory aims are for descriptive purposes and p-values resulting from these analyses will not be evaluated for statistical significance.

1. **Secondary urinary outcomes:** To compare treatment with Botox A to treatment with mid-urethral sling for improving the number of urinary incontinence episodes on bladder diary after 12 months.

   Changes from baseline in bladder diary outcomes will be calculated and analyzed using the methods described for the analysis of the primary outcome. For urinary frequency, women reporting on average >8 voids/24 hours at baseline will be considered symptomatic, and normalization of voiding frequency will be defined as $\leq$ 8 voids/24 hours at 6 months. A 50% improvement will be defined as a reduction by half in the number of voids that patients had at baseline. The number of women who had normalization of voiding frequency and 50% improvement at 6 months will be compared between groups separately and collectively. We will also assess the proportion of women who had worsening of urinary frequency (includes women who developed de novo frequency and those who worsened). These dichotomous outcomes will be

analyzed using logistic regression, controlling for the design effects of stratification by site and age group.

2. **Predictors of poor treatment response:** To develop models to identify baseline predictors of change of MUI, OAB, and SUI outcomes measured using the UDI between baseline and 12 months post-treatment

Regression models will be created to identify predictors of change from baseline to 12 months for UDI total score and stress and irritative subscale scores. Potential predictors will include age, diary parameters such as number of UUI episodes/3 days, functional bladder capacity, stress and irritative bother severity at baseline, type of urinary incontinence (stress-predominant or urgency-predominant: see definition below), and co-existing anterior vaginal wall prolapse. The relationship between potential predictors and outcomes will be explored in models that include one predictor plus stratification factors (site and age group). Predictive models will be constructed using backward selection of predictors. The impact of collinearity between predictors will be assessed and the final model modified as necessary.

The definition of the type of urinary incontinence, stress- or urgency predominant will be based on 2 questions of the Urogenital Distress Inventory,

1. "Do you experience urine leakage associated with a feeling of urgency?"
2. "Do you experience urine leakage related to physical activity, coughing, or sneezing?"

Based on our inclusion criteria, patients must report at least "moderate" bother on both questions.
If the subject reports greater bother on the stress question, she will be classified as stress-predominant MUI. If the subject reports greater bother on the urgency question, she will be classified as urgency-predominant MUI. If the subject reports equal bother on both questions, she will be classified as "balanced" MUI.

3. **Quality of life and global impression:** To compare quality of life outcomes and Patient Global Impression-Improvement (PGI-I), Patient Global Impression-Severity (PGI-S) between groups randomized to Botox versus sling after 12 months.

   Changes from baseline in all quality of life PRO instruments and global impression will be calculated and analyzed using the methods described for the analysis of the primary outcome

4. **Safety and additional treatments:** To describe rates of reoperation (sling revision) after mid-urethral sling and intermittent catheterization due to voiding dysfunction/partial urinary retention after Botox injection, to compare the proportion of women in each group with UTI and recurrent UTI, rates of other serious and non-serious adverse events, and to compare the proportion of women in each group initiating additional (off protocol) treatment for SUI and/or OAB, and the types of additional treatment (off protocol Botox or mid-urethral sling, behavioral therapy, medications, other).

   Groups will be summarized with n and percent, and rates will be compared with chi-square tests.

5. **Cost-effectiveness analysis:** To determine the cost effectiveness of Botox A injection versus mid-urethral sling for the treatment of MUI symptoms at 12 months.

   Differential mean costs and differential mean QALYs between the two treatment groups will be estimated using multiple regression analysis. Specifically, a generalized linear model with appropriate link function (e.g., log-link) and response probability distribution (e.g., gamma

distribution) will be used to analyze costs due to the potential skewness and heteroscedasticity of medical expenditure data, while an ordinary least squares regression will be used for analyzing QALY data. The models will account for stratification factors of study site and age group, as well as other characteristics of the subjects that are found to differ significantly between the groups. When estimating QALYs, we will also adjust for subjects' baseline utility scores to account for potential imbalance in baseline utility between the two treatment groups.

We will calculate the incremental cost-effectiveness ratio (ICER), which is the differential mean costs divided by the differential mean QALYs between the two groups, to assess the additional costs associated with each additional QALY gained. Our base case analysis will be conducted based on subjects with complete data. Sensitivity analysis will be conducted to include subjects with incomplete data using the multiple imputation method. Non-parametric bootstrapping resampling technique will be used to derive the 95% confidence interval for the ICER. In addition, cost-effectiveness acceptability curve (CEAC) will be generated to illustrate the likelihood that one treatment is more cost-effective than the other with various ceiling cost-effectiveness ratios.

We plan to conduct supplemental analyses using alternative outcome measures, such as incremental cost per reduction in UIE/day.

The cost-effectiveness evaluations will be conducted as within-trial comparisons. A decision analytic model will also be developed from trial data to evaluate the trajectory of the cost-effectiveness ratio.

6. **UDI MID:** To explore MIDs for UDI total score and stress and irritative subscores for this MUI population

MIDs will be calculated using anchor- and distribution-based approaches. Potential anchors include the PGSC, global impression of change, and incontinence episodes from the bladder diary.