Back Pain Consortium Research Program (BACPAC)

The BEST Trial: Biomarkers for Evaluating Spine Treatments (BEST)

ClinicalTrials.gov Statistical Analysis Plan (SAP)

SAP Version: 1.0
SAP Version Date: 05/19/2025

Protocol Number: 00057948
Protocol Version 7.0

National Clinical Trial (NCT) Number: NCT05396014

Sponsor: The Helping to End Addiction Long-term℠ Initiative
(NIH HEAL Initiative℠)

Funded by: National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)

SAP Approval Page

Study Title: The BEST Trial: Biomarkers for Evaluating Spine Treatments (BEST)
SAP Version: 1.0
SAP Version Date: 05/19/2025

We, the undersigned, have read and approved of this SAP and agree on its content.

_____
BEST Protocol Co-Chair – Daniel Clauw, MD

_____
BEST Protocol Co-Chair – Gwendolyn Sowa, MD, PhD

_____
BEST Protocol Co-Chair – Matthew Mauck, MD, PhD

_____
BEST Protocol Co-Chair – Kevin Anstrom, PhD

_____
BEST Protocol Statistician – Anastasia Ivanova, PhD

_____
BEST Protocol Statistician – Bryce Rowland, PhD

Introduction

Background and Rationale

Chronic low-back pain, defined as pain lasting three months or more with pain occurring on most days, affects 10-20% of adults in the United States. Those facing chronic low-back pain face physical pain, limited mobility, and mental health symptoms. Chronic low-back pain is second only to cancer in the number of opioids prescribed, and opioid are the most commonly prescribed class of drugs for low-back pain, despite limited evidence of efficacy in chronic pain and the significant risks of side effects, addiction, and death. Current treatments, covering a broad range, do not cure chronic low-back pain for most patients.[1]

Through the NIH Helping to End Addiction Long-term℠ Initiative, the NIH seeks an improved understanding of the underlying biological mechanism of chronic pain and supports the discovery and testing of novel non-addictive pain treatments to stem the ongoing opioid crisis and support the translation of scientific findings into clinical practice.

While a broad range of treatments exists, current treatments do not adequately resolve chronic low-back pain for most patients. Systematic reviews of common treatments find small to moderate evidence to support many currently used treatments. BACPAC, a funded component of the HEAL Initiative℠, is a multisite consortium that will advance knowledge of the etiology and treatment of chronic low-back pain by developing an understanding of the mechanisms contributing to chronic low-back pain and identifying specific treatments or combinations of treatments that are most effective in identifiable subgroups of participants. The BEST Trial (Biomarkers for Evaluating Spine Treatments) is a NIAMS-sponsored clinical trial conducted by the BACPAC research program. The multi-site, sequential, multiple assignment randomized trial (SMART) was designed to evaluate four evidence-based interventions cLBP and inform a precision medicine approach to treating cLBP.

Objectives

The BEST trial was powered for the study's primary objective of estimating an algorithm to assign sequences of two cLBP treatments based on phenotypic markers and an individual patient's response to the initial treatment (i.e., a dynamic treatment regime (DTR)) that optimizes effectiveness, as characterized by the primary and secondary endpoints. In this outcome analysis, we will report the ITT-least squares adjusted mean of the 24-week outcomes of the BEST Trial and associated 95% confidence intervals. It is important to reiterate that the BEST trial was powered for biomarker discovery. This analysis describes the primary and secondary outcomes of the study.

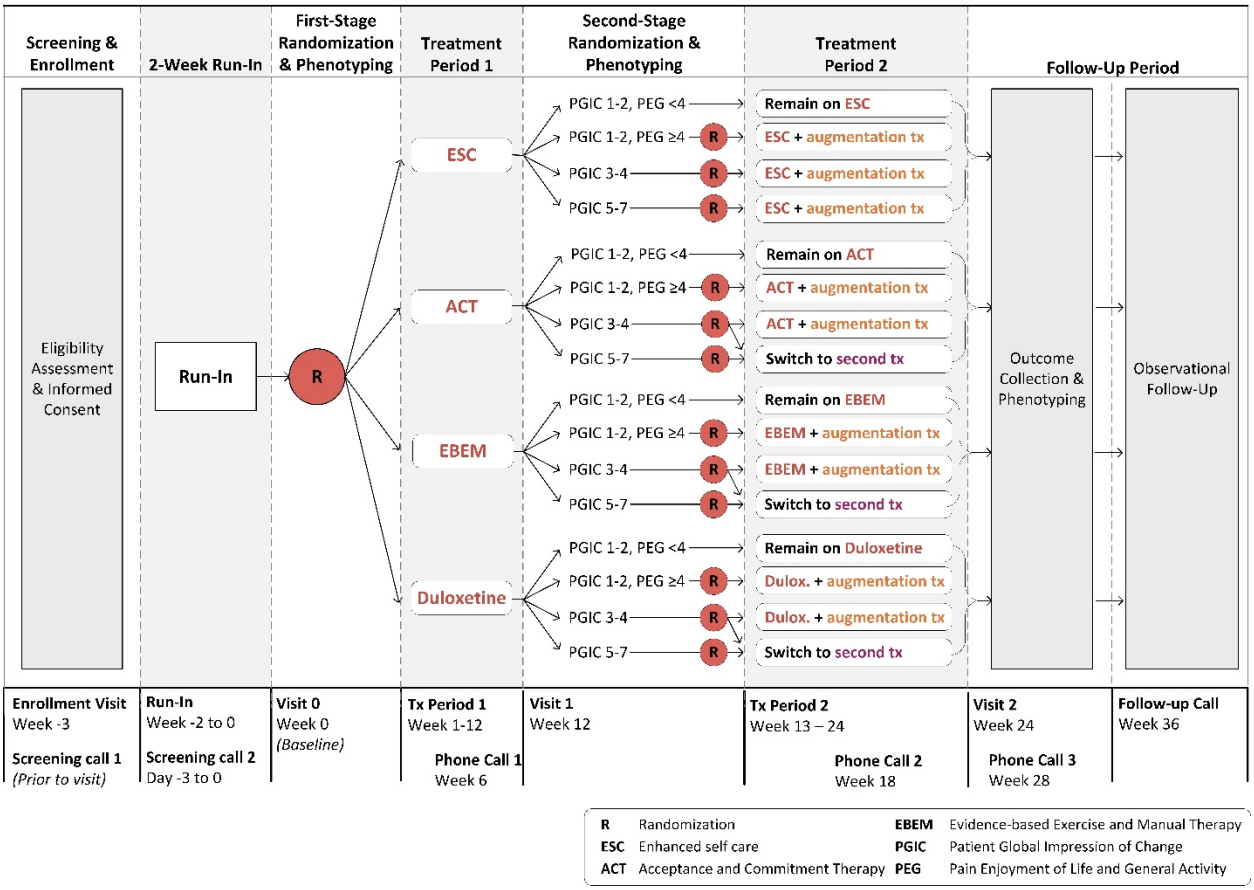
In this analysis, we aim to:

- Report on BEST Trial study flow, baseline characteristics, and adverse events
- Estimate 24-week least squares means for study outcomes by Stage 1 treatment assignment.

Trial/Study Design

Study data for this analysis will be from the BEST Trial, which has been described elsewhere.[1-3] The trial follows a two stage sequential multiple assignment random trial SMART trial plan.[2] The trial is split into two 12-week intervention stages with 3 study visits that occurred at the beginning of Stage 1 (Visit 0/Week 0), Stage 2 (Visit 1/Week 12), and the end of Stage 2(Visit 2/Week 24). At baseline (Visit 0/Week 0) participants were randomized to one of four interventions. To participate in the study participants were required to be eligible for a minimum of 3 of the treatments at baseline. A re-randomization happened at Week 12, where participants were assigned to one of four responder status categories using a combination of Week 12 Pain, Enjoyment of Life, and General Activity (PEG) and Patient Global Impression of Change (PGIC) scores. Participants with Week 12 PGIC 1-2 and PEG <= 4 maintained their initial treatment, participant with Week 12 PGIC 1-2 and PEG > 4 were randomized to augment a second treatment, participants with a Week 12 PGIC 3-4 were randomized to either augment a second treatment or switch treatments, and participants with a Week 12 PGIC of 5-7 or participants who could not tolerate their initial treatment were switched to a new treatment. Randomization at the baseline visit was stratified by years of cLBP, depression/anxiety, consent to optional phenotyping, and current use of opioids. The details of the trial are illustrated in the schema below.

Schema



Sample Size
The sample size calculation for the BEST Trial is based on the ability to accurately estimate the value of the estimated optimal dynamic treatment regime relative to an oracle estimator and is described in the study protocol.

Statistical Principles

Confidence Levels and p-values
Level of statistical significance

Our results will not include any hypothesis testing. All confidence intervals for least squares mean outcome estimates will be reported at the 95% confidence level.

Multiplicity Adjustments

We will not use a multiplicity adjustment to adjust for the width of the confidence interval because the least squares mean is not a primary estimand of interest in the trial's statistical design. The primary estimand of interest is the value of the estimated optimal dynamic treatment regime, which is outside the scope of the outcome analysis presented here.

Adherence and Protocol Deviations
The analysis will summarize results from all randomized participants. Adherence to interventions will not be considered in this analysis.

Analysis Populations
The population used for statistical analysis consists of the 805 randomized participants to a Stage 1 treatment.

Trial Population

Eligibility
To be eligible for participation in the BEST Trial an individual must meet all of the following inclusion criteria:

- Ability to read and understand English
- Provision of signed and dated informed consent form(s)
- Willing and able to receive study-related messages and survey links via email
- Willing and able to receive study-related phone calls
- Age 18 years old or older
- Low-back pain for at least 3 months and occurring on at least half the days in the past 6 months
- Contraindicated to no more than one of the study interventions at the time of eligibility
- assessment(s)
- Eligible to receive at least three of the four study interventions and willing to receive any
- intervention for which they are eligible
- A PEG score 4 or higher prior to the Run-in period
- Willing and able to undergo required phenotyping as defined in Section 9

- Regular reliable access to an internet-enabled device such as a smart phone, tablet, or laptop
- computer
- Meet Run-in period engagement eligibility criteria:
- Completion of two Run-in study information modules prior to period 1 randomization
- (Visit 0)
- Low-back pain more severe than pain in other parts of the body
- Available to complete the full study protocol (approximately 9 months

An individual who meets any of the following criteria will be excluded from participation in this study:

- Pregnant at the time of Visit 0 (Baseline)
- Affirmative participant response to any of the following conditions:
    - Progressive neurodegenerative disease
    - History of discitis osteomyelitis (spine infection) or spine tumor
    - BEST Trial Protocol v7.0 20 9/1/2023
    - History of ankylosing spondylitis, rheumatoid arthritis, polymyalgia rheumatica, psoriatic
    - arthritis, or lupus
    - History of cauda equina syndrome or spinal radiculopathy with functional motor deficit
    - (strength <4/5 on manual motor testing)
    - Diagnosis of any vertebral fracture in the last 6 months
    - Osteoporosis requiring pharmacologic treatment other than vitamin D, calcium
    - supplements, or bisphosphonates.
    - History of any bone-related cancer or cancer that metastasized to the bone
    - Currently in treatment for any non-skin cancer or plan to start non-skin cancer
    - treatment in the next 12 months
    - History of any non-skin cancer treatment in the last 24 months
    - Visual or hearing difficulties that would preclude participation
    - Uncontrolled drug/alcohol addiction
    - Individuals actively pursuing disability or workers compensation or involved in active
    - personal injury-related litigation
    - Currently participating in another interventional pain study
- Any condition that, in the opinion of the investigator, would preclude the patient from being able to safely participate in in the trial

Recruitment

Recruitment for the BEST Trial has been described extensively in the trial protocol and protocol paper.[1]

Withdrawal/follow-up

Study flow reporting is specified in Appendix 1.

Timing of withdrawal/lost to follow-up data

Lost to follow-up status was determined at the subsequent study visit at 12-week intervals. Study withdrawal was determined as soon as a participant withdrew consent.

Baseline Characteristics

List of Baseline Characteristics to be summarized

The baseline characteristics analyzed are:

- Age
- Sex Assigned at Birth
- Gender Identity
- Ethnicity
- Race
- Highest Level of Education Completed
- Current Employment Status
- Current Relationship Status
- Number of People Living in Household
- Annual household income from all sources
- PEG score (0 to 10)
- Self-Reported low-back pain duration (months)
- Self-reported low-back pain duration (category)
- Low-back Pain Frequency in the Past 6 months
- Low-back Pain Specific Pain Intensity
- Ever had a low back operation?
- When was the last back operation?
- Any back operations involve a spinal fusion?
- Ever unemployed for 1 or more months due to low back pain
- Ever file or awarded worker's compensation claim related to back problem
- Involved in a lawsuit or legal claim related to back problem
- Ever applied for, or received, disability insurance for pain condition
- BMI
- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg)
- Heart Rate (bpm)
- Ever had Hip Replacement surgery?
- Ever had Knee Replacement surgery?
- Observed Gait
- Tobacco Use in Previous 12 Months
- Alcohol Use in Previous 12 Months
- Drug Use in Previous 12 Months
- Prescription Drug Used Not as Intended in Previous 12 Months
- Currently Taking Opioid Medication (daily)
- Previously diagnosed with COVID-19
- Long COVID
- COVID-19 Vaccine

- Fear avoidance beliefs about physical activity raw scoring scale 2
- GAD-2 Raw Score
- Keele STarT Back Screening Tool Risk
- Oswestry Disability Index Percentage
- Pain Catastrophizing Scale SF-6
- Pain Detect Questionnaire Raw Score
- PHQ-2 Raw Score
- HEAL Positive Outlook Raw Score
- PROMIS-Cognitive Function – Abilities 2a Raw Score
- General Sensory Sensitivity Score – External
- General Sensory Sensitivity Score – Interoception
- General Sensory Sensitivity Score – Total
- Self-reported Sleep Duration in the Past Month (hours)
- Symptom Severity Index
- Widespread Pain Raw Score
- Stomach Pain in the Past 4 Weeks
- Headaches in the Past 4 Weeks
- Radiating Pain to Buttock/Thigh, Past 2 Weeks
- Radiating Pain to Below Knee, Past 2 Weeks
- Social Determinants of Health: Transportation Needs
- Social Determinants of Health: Healthcare Needs
- Social Determinants of Health: Food Insecurity
- Social Determinants of Health: Food Money
- Social Determinants of Health: Utilities
- Social Determinants of Health: Stable Housing
- Social Determinants of Health: Emotional Support
- Social Determinants of Health: Number of Close Friends
- Perceived Discrimination: Race/Ethnicity
- Perceived Discrimination: Orientation/Gender Identity

Details of how baseline characteristics will be summarized

The baseline characteristics are reported as in Appendix 2. Continuous variables are summarized with means and standard deviations. Categorical variables are analyzed with counts and percentages. More details and analysis of the baseline population are described elsewhere.[2]

Analysis

Analysis Methods

A linear mixed-effect model (LMM) was used to model the least squares mean of the primary and secondary outcomes at Visit 2 (Week 24) by the Stage 1 treatment assignment for all continuous outcomes. Random intercepts and a linear random slope for time were included to account for subject-level variability. The randomization and minimization algorithm variables are included as

fixed effects. Week indicators as fixed effects were included to model non-linear effects of time. An additional fixed effect was included for the number of treatments a participant was eligible for at the baseline visit. Estimation was conducted using REML and modern optimization approaches with the lme4 package in R. Least squares means were estimated for each Week 24 outcome, along with associated 95% confidence intervals, using the emmeans package in R. Due to the small number of participants on opioids in the BEST Trial sample, we only report counts of participants self-reporting opioid use at Week 24.

Outcome Definitions

Outcome Name	Outcome Type (primary/secondary)	Specification of outcome and timing	Specific measurement and units	Calculation to derive outcome	Analysis Method
Patient-Reported Pain Intensity and Interference Score	Primary	PEG Score measured at the Visit 2 (Week 24)	The PEG scale is based on a participant rating of their pain in 3 areas: pain intensity, pain interference with enjoyment of life, and pain interference with general activities. A rating is given ranging from -10 (lowest) to 10 (highest) on each of these scales. We use the average of these ratings, with higher scores indicating increased pain intensity and interference at 24 Weeks compared to Baseline.	Calculate the PEG Score by averaging the responses from each of the 3 scales.	LMM
Pain Interference	Secondary	Pain Interference	Ratings of the consequences	Calculate the raw score by	LMM

		at Visit 2 (Week 24)	of pain on relevant aspects of one's life. This includes the extent to which the pain hinders engagement with social, cognitive, emotional, physical, and recreational activities. Results range from -34 to 34 (t-scores range from 41.6 to 75.6), with higher scores indicating increased pain interference at 24 Weeks compared to Baseline.	summing all appropriate responses, then translate the raw score to the corresponding T-Score	
Physical Function	Secondary	Physical Function at Visit 2 (Week 24)	Measured by the PROMIS-PF Short Form 6b. The PROMIS-PF Short Form 6b is a series of 6 questions. Results range from -37.1 to 37.1 (t-scores range from 21.6 to 58.7), with higher scores indicating increased physical functioning at 24 Weeks	Calculate the raw score by summing all appropriate responses, then translate the raw score to the corresponding T-Score	LMM

			compared to Baseline.		
Depression Score	Secondary	Depression at Visit 2 (Week 24)	Measured by the PROMIS 4-item depression scale from the PROMIS 29 profile. The PROMIS 4-item depression scale from the PROMIS 29 profile is a series of 4 questions. A rating of negative mood, views of self, as well as decreased positive affect and engagement. Results range from -38.4 to 38.4 (t-scores range from 41.0 to 79.4), with higher scores indicating increased depression at 24 Weeks compared to Baseline.	Calculate the raw score by summing all appropriate responses, then translate the raw score to the corresponding T-Score	LMM
Anxiety Score	Secondary	Anxiety at Visit 2 (Week 24)	Measured by the PROMIS Emotional Distress-Anxiety scale (PROMIS-EDA 4a). The PROMIS-EDA 4a is a series of 4 questions. A rating of fear, anxious	Calculate the raw score by summing all appropriate responses, then translate the raw score to the corresponding T-Score	LMM

			misery, hyperarousal, and somatic symptoms related to arousal. Results range from -41.3 to 41.3 (t-scores range from 40.3 to 81.6), with higher scores indicating increased anxiety at 24 Weeks compared to Baseline.		
Sleep Disturbance	Secondary	Sleep Disturbance at Visit 2 (week 24)	Measured by the PROMIS short form 6a. The PROMIS short form 6a is a series of 6 questions. A rating of sleep quality, sleep depth, and restoration associated with sleep. Results range from -44.4 to 44.4 (t-scores range from 31.7 to 76.1), with higher scores indicating increased sleep disturbance at 24 Weeks compared to Baseline.	Calculate the raw score by summing all appropriate responses, then translate the raw score to the corresponding T-Score	LMM
Sleep Duration	Secondary	Response to BACPAC sleep disturbance	Response to "During the past month, how many		LMM

		question at Visit 2 (Week 24)	hours and minutes of actual sleep did you get at night?" measured in Hours and Minutes. Results range from -24 to 24 hours, with higher number of hours indicating increased sleep duration at 24 Weeks compared to Baseline.		
Opioid Incidence	Secondary	Opioid incidence reported at Visit 2 (Week 24)	Self-reported Yes/No/Not Sure response to a single question, "Are you currently taking any opioid pain medication on a daily basis?"		Counts

Missing Data

We assume the data are missing at random. We include all subjects randomized at baseline in the analysis, regardless of the number of future observations. There is no missingness in model covariates because they were required for randomization at the baseline visit.

Additional Analyses

Additional reporting is included in this submission for adverse events (Appendix 3)

References

1. Mauck MC, Barth KS, Bell KM, et al. The Design and Rationale of the Biomarkers for Evaluating Spine Treatments (BEST) Trial: A Sequential Multiple Assignment Randomized Trial. *Pain Med*. Apr 9 2025; doi:10.1093/pm/pnaf032.

2. Rowland B, Barth KS, Bell KM, et al. *Baseline Characteristics of Participants in the BEST Clinical Trial: A Sequential Multiple Assignment Randomized Trial for Chronic Low Back Pain*. 2025.

3. Sperger J, Kidwell KM, Mauck MC, et al. Statistical Design and Rationale of the Biomarkers for Evaluating Spine Treatments (BEST) Trial. Preprint on arxiv.org, 2025.

Table 1.0. BEST Trial Recruitment Flow Diagram

All Randomized Participants at Baseline

Arm/Group Title	Acceptance and Commitment Therapy (ACT) Alone	ACT Plus Duloxetine, Moderate Treatment Response	ACT Plus Evidence Based Exercise and Manual Therapy (EBEM), Moderate Stage 1 Treatment Response	ACT Plus Enhanced Self-Care Therapy (ESC), Moderate Stage 1 Treatment Response	ACT Plus Duloxetine, Low Stage 1 Treatment Response	ACT Plus Evidence Based Exercise and Manual Therapy (EBEM), Low Stage 1 Treatment Response	ACT Plus Enhanced Self-Care Therapy (ESC), Low Stage 1 Treatment Response
Arm/Group Description	This arm included both participants who had a strong response to ACT in Stage 1 and participants who were not assessed for response at the end of Stage 1. In Stage 1, participants randomized to ACT took part in 12 therapy sessions over the course of 12 weeks. Each participant was scheduled for a combination of 4 remote face-to-face visits and 8 therapist-supported online sessions, which included a self-directed module combined with personalized provider coaching delivered via an online messaging system. In Stage 2, strong responders were encouraged to continue to practice ACT skills at home without therapist supervision and were given access to an additional 11 ACT audio modules.	This arm included participants who had a moderate response to ACT in Stage 1 and started Duloxetine in Stage 2 while continuing ACT treatment.	This arm included participants who had a moderate response to ACT in Stage 1 and started EBEM treatment in Stage 2 while continuing ACT.	This arm included participants who had a moderate response to ACT in Stage 1 and started ESC treatment in Stage 2 while continuing ACT.	This arm included participants who had a low response to ACT in Stage 1 and started Duloxetine treatment in Stage 2 while continuing ACT.	This arm included participants who had a low response to ACT in Stage 1 and started EBEM treatment in Stage 2 while continuing ACT.	This arm included participants who had a low response to ACT in Stage 1 and started ESC treatment in Stage 2 while continuing ACT.
Stage 1 Intervention (Weeks 0-12)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Screening (Week 12 Pre-Randomization)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Intervention (Weeks 12-24)							
Started							
Completed							
Not Completed							
Reason Not Completed							

Arm/Group Title	ACT Followed by Duloxetine, Low Stage 1 Treatment Response	ACT Followed by Evidence Based Exercise and Manual Therapy (EBEM), Low Stage 1 Treatment Response	ACT Followed by Enhanced Self-Care Therapy (ESC), Low Stage 1 Treatment Response	ACT Followed by Duloxetine, Poor Stage 1 Treatment Response	ACT Followed by Evidence Based Exercise and Manual Therapy (EBEM), Poor Stage 1 Treatment Response	ACT Followed by Enhanced Self-Care Therapy (ESC), Poor Stage 1 Treatment Response	Duloxetine Alone
Arm/Group Description	This arm included participants who had a low response to ACT in Stage 1 and started Duloxetine treatment in Stage 2. ACT treatment was discontinued.	This arm included participants who had a low response to ACT in Stage 1 and started EBEM treatment in Stage 2. ACT treatment was discontinued.	This arm included participants who had a low response to ACT in Stage 1 and started ESC treatment in Stage 2. ACT treatment was discontinued.	This arm included participants who had a poor response to ACT in Stage 1 and started Duloxetine treatment in Stage 2. ACT treatment was discontinued.	This arm included participants who had a poor response to ACT in Stage 1 and started EBEM treatment in Stage 2. ACT treatment was discontinued.	This arm included participants who had a poor response to ACT in Stage 1 and started ESC treatment in Stage 2. ACT treatment was discontinued.	This arm included both participants who had a strong response to Duloxetine in Stage 1 and participants who were not assessed for response at the end of Stage 1.
Stage 1 Intervention (Weeks 0-12)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Screening (Week 12 Pre-Randomization)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Intervention (Weeks 12-24)							
Started							
Completed							
Not Completed							
Reason Not Completed							

Arm/Group Title	Duloxetine Plus Acceptance and Commitment Therapy (ACT), Moderate Stage 1 Treatment Response	Duloxetine Plus Evidence Based Exercise and Manual Therapy (EBEM), Moderate Stage 1 Treatment Response	Duloxetine Plus Enhanced Self-Care Therapy (ESC), Moderate Stage 1 Treatment Response	Duloxetine Plus Acceptance and Commitment Therapy (ACT), Low Stage 1 Treatment Response	Duloxetine Plus Evidence Based Exercise and Manual Therapy (EBEM), Low Stage 1 Treatment Response	Duloxetine Plus Enhanced Self-Care Therapy (ESC), Low Stage 1 Treatment Response	Duloxetine Followed by Acceptance and Commitment Therapy (ACT), Low Stage 1 Treatment Response
Arm/Group Description	This arm included participants who had a moderate response to Duloxetine in Stage 1 and started ACT treatment in Stage 2 while continuing Duloxetine.	This arm included participants who had a moderate response to Duloxetine in Stage 1 and started EBEM treatment in Stage 2 while continuing Duloxetine.	This arm included participants who had a moderate response to Duloxetine in Stage 1 and started ESC treatment in Stage 2 while continuing Duloxetine.	This arm included participants who had a low response to Duloxetine in Stage 1 and started ACT treatment in Stage 2 while continuing Duloxetine.	This arm included participants who had a low response to Duloxetine in Stage 1 and started EBEM treatment in Stage 2 while continuing Duloxetine.	This arm included participants who had a low response to Duloxetine in Stage 1 and started ESC treatment in Stage 2 while continuing Duloxetine.	This arm included participants who had a low response to Duloxetine in Stage 1 and started ACT treatment in Stage 2. Duloxetine treatment was discontinued.
Stage 1 Intervention (Weeks 0-12)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Screening (Week 12 Pre-Randomization)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Intervention (Weeks 12-24)							
Started							
Completed							
Not Completed							
Reason Not Completed							

Arm/Group Title	Duloxetine Followed by Evidence Based Exercise and Manual Therapy (EBEM), Low Stage 1 Treatment Response	Duloxetine Followed by Enhanced Self-Care Therapy (ESC), Low Stage 1 Treatment Response	Duloxetine Followed by Acceptance and Commitment Therapy (ACT), Poor Stage 1 Treatment Response	Duloxetine Followed by Evidence Based Exercise and Manual Therapy (EBEM), Poor Stage 1 Treatment Response	Duloxetine Followed by Enhanced Self-Care Therapy (ESC), Poor Stage 1 Treatment Response	Evidence Based Exercise and Manual Therapy (EBEM) Alone	EBEM Plus Duloxetine, Moderate Stage 1 Treatment Response
Arm/Group Description	This arm included participants who had a low response to Duloxetine in Stage 1 and started EBEM treatment in Stage 2. Duloxetine treatment was discontinued.	This arm included participants who had a low response to Duloxetine in Stage 1 and started ESC treatment in Stage 2. Duloxetine treatment was discontinued.	This arm included participants who had a poor response to Duloxetine in Stage 1 and started ACT treatment in Stage 2. Duloxetine treatment was discontinued.	This arm included participants who had a poor response to Duloxetine in Stage 1 and started EBEM treatment in Stage 2. Duloxetine treatment was discontinued.	This arm included participants who had a poor response to Duloxetine in Stage 1 and started ESC treatment in Stage 2. Duloxetine treatment was discontinued.	This arm included both participants who had a strong response to EBEM in Stage 1 and participants who were not assessed for response at the end of Stage 1.	This arm included participants who had a moderate response to EBEM in Stage 1 and started Duloxetine treatment in Stage 2 while continuing EBEM.
Stage 1 Intervention (Weeks 0-12)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Screening (Week 12 Pre-Randomization)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Intervention (Weeks 12-24)							
Started							
Completed							
Not Completed							
Reason Not Completed							

Arm/Group Title	EBEM Plus Acceptance and Commitment Therapy (ACT), Moderate Stage 1 Treatment Response	EBEM Plus Enhanced Self-Care Therapy (ESC), Moderate Stage 1 Treatment Response	EBEM Plus Duloxetine, Low Stage 1 Treatment Response	EBEM Plus Acceptance and Commitment Therapy (ACT), Low Stage 1 Treatment Response	EBEM Plus Enhanced Self-Care Therapy (ESC), Low Stage 1 Treatment Response	EBEM Followed by Duloxetine, Low Stage 1 Treatment Response	EBEM Followed by Acceptance and Commitment Therapy (ACT), Low Stage 1 Treatment Response
Arm/Group Description	This arm included participants who had a moderate response to EBEM in Stage 1 and started ACT treatment in Stage 2 while continuing EBEM.	This arm included participants who had a moderate response to EBEM in Stage 1 and started ESC treatment in Stage 2 while continuing EBEM.	This arm included participants who had a low response to EBEM in Stage 1 and started Duloxetine treatment in Stage 2 while continuing EBEM.	This arm included participants who had a low response to EBEM in Stage 1 and started ACT treatment in Stage 2 while continuing EBEM.	This arm included participants who had a low response to EBEM in Stage 1 and started ESC treatment in Stage 2 while continuing EBEM.	This arm included participants who had a low response to EBEM in Stage 1 and started Duloxetine treatment in Stage 2. EBEM treatment was discontinued.	This arm included participants who had a low response to EBEM in Stage 1 and started ACT treatment in Stage 2. EBEM treatment was discontinued.
Stage 1 Intervention (Weeks 0-12)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Screening (Week 12 Pre-Randomization)							
Started							
Completed							
Not Completed							
Reason Not Completed							
Stage 2 Intervention (Weeks 12-24)							
Started							
Completed							
Not Completed							
Reason Not Completed							

Arm/Group Title	EBEM Followed by Enhanced Self-Care Therapy (ESC), Low Stage 1 Treatment Response	EBEM Followed by Duloxetine, Poor Stage 1 Treatment Response	EBEM Followed by Acceptance and Commitment Therapy (ACT), Poor Stage 1 Treatment Response	EBEM Followed by Enhanced Self-Care Therapy (ESC), Poor Stage 1 Treatment Response	Enhanced Self-Care Therapy (ESC) Alone	ESC Plus Duloxetine, Moderate to Low Stage 1 Treatment Response	ESC Plus Acceptance and Commitment Therapy (ACT), Moderate to Low Stage 1 Treatment Response	ESC Plus Evidence Based Exercise and Manual Therapy (EBEM), Moderate to Low Stage 1 Treatment Response
Arm/Group Description	This arm included participants who had a low response to ESC in Stage 1 and started Duloxetine treatment in Stage 2. EBEM treatment was discontinued.	This arm included participants who had a poor response to EBEM in Stage 1 and started Duloxetine treatment in Stage 2. EBEM treatment was discontinued.	This arm included participants who had a poor response to EBEM in Stage 1 and started ACT treatment in Stage 2. EBEM treatment was discontinued.	This arm included participants who had a poor response to EBEM in Stage 1 and started ESC treatment in Stage 2. EBEM treatment was discontinued.	This arm included both participants who had a strong response to ESC in Stage 1 and participants who were not assessed for response at the end of Stage 1.	This arm included participants who had a moderate to poor response to ESC in Stage 1 and started Duloxetine treatment in Stage 2. ESC treatment was discontinued.	This arm included participants who had a moderate to poor response to ESC in Stage 1 and started ACT treatment in Stage 2. ESC treatment was discontinued.	This arm included participants who had a moderate to poor response to ESC in Stage 1 and started EBEM treatment in Stage 2. ESC treatment was discontinued.
Stage 1 Intervention (Weeks 0-12)								
Started								
Completed								
Not Completed								
Reason Not Completed								
Stage 2 Screening (Week 12 Pre-Randomization)								
Started								
Completed								
Not Completed								
Reason Not Completed								
Stage 2 Intervention (Weeks 12-24)								
Started								
Completed								
Not Completed								
Reason Not Completed								

Table 1: Tabular Characteristics of Randomized Population

Variable[1]	ACT (N=203)	Duloxetine (N=198)	EBEM (N=199)	ESC (N=205)	Total (N=805)
Age					
Age Group					
< 18 years					
18-65 years					
>65 years					
Sex Assigned at Birth, n(%)[2]					
Female					
Male					
Missing					
Gender Identity, n(%)					
Female					
Male					
Non-Binary					
Unknown					
Missing					
Ethnicity, n(%)					
Hispanic or Latino					
Not Hispanic or Latino					
Unknown/Not Reported					
Missing					
Race, n(%)					
White					
Black, African American, or Black Multiracial					
Asian or Asian Multiracial					
Indigenous or Indigenous Multiracial					
Unknown or Not Reported					
Missing					
Highest Level of Education Completed, n(%)					
Did not complete Secondary School or High School					
High School or Secondary School Degree Complete					

Associate's or Technical Degree Complete					
College or Baccalaureate Degree Complete					
Doctoral or Postgraduate Education					
Missing					
Current Employment Status, n(%)					
Full-time employment					
Not employed[3]					
Part-time employment					
Missing					
Current Relationship Status, n(%)					
Married					
Never married					
Domestic Partner					
Widowed					
Divorced or Separated					
Missing					
Number of People Living in Household, n(%)					
1					
2					
3					
4					
5					
6					
7					
8					
Missing					
Annual household income from all sources, n(%)					
Less than $10,000					
$10,000 to $24,999					
$25,000 to $34,999					

$35,000 to $49,999					
$50,000 to $74,999					
$75,000 to $99,999					
$100,000 to $149,999					
$150,000 to $199,999					
$200,000 or more					
Prefer not to answer					
Missing					
PEG score (0 to 10)					
PEG score, median (range)					
Self-reported low-back pain duration (months)					
Self-reported low-back pain duration, n(%)					
< 1 year					
1-5 years					
> 5 years					
Missing					
Low Back Pain Frequency in the Past 6 months, n(%)					
Every day or nearly every day					
At least half the days					
Low Back Pain Specific Pain Intensity, (0 to 10)					
Low Back Pain SPI, median (range)					
Ever had low back operation, n(%)					
No					
Yes, at least one					
Missing					
When was last back operation, n(%)[4]					
Less than 6 months ago					
More than 6 months ago but less than 1 year ago					

Between 1 and 2 years ago					
More than 2 years ago					
Any back operations involve a spinal fusion, n(%)[4]					
No					
Yes					
Not Sure					
Ever unemployed for 1 or more months due to low back pain, n(%)					
No					
Yes					
Does not apply					
Missing					
Ever filed or awarded worker's compensation claim related to back problem, n(%)					
No					
Yes					
Does not apply					
Missing					
Involved in a lawsuit or legal claim related to back problem, n(%)					
No					
Yes					
Not sure					
Missing					
Ever applied for, or received, disability insurance for pain condition, n(%)					
No					
Yes					
Missing					
BMI					
Systolic Blood Pressure (mmHg)					
Diastolic Blood Pressure (mmHg)					
Heart Rate (bpm)					

Ever Hip Replacement Surgery					
Yes					
No					
Missing					
Ever Knee Replacement Surgery					
Yes					
No					
Missing					
Observed Gait					
Normal					
Antalgic					
Missing					
Tobacco Use in Previous 12 Months					
Daily or Almost Daily					
Less Than Daily					
Never					
Missing					
Alcohol Use in Previous 12 Months[5]					
At Least Weekly					
Monthly					
Less Than Monthly					
Never					
Missing					
Drug Use in Previous 12 Months					
Daily or Almost Daily					
Between Monthly to Weekly					
Less Than Monthly					
Never					
Missing					
Prescription Drug Used Not as Intended in Previous 12 Months					
At Least Once					
Never					
Missing					

Currently Taking Opioid Medication (daily), n(%)					
Yes					
No					
Not sure					
Previously Diagnosed with COVID-19					
Yes					
No					
Not Sure					
Prefer not to Answer					
Missing					
Long COVID					
Yes					
No					
Not Sure					
Prefer not to Answer					
No self-reported COVID-19 Diagnosis					
COVID-19 Vaccine					
Yes					
No					
Not Sure					
Prefer Not to Answer					
Fear-avoidance beliefs about physical activity raw scoring scale 2					
GAD-2 Raw Score					
Keele STarT Back Screening Tool Risk, n(%)					
Low Risk					
Medium Risk					
High Risk					
Missing					
Oswestry Disability Index Percentage					
ODI Percentage, median (range)					
Pain Catastrophizing Scale SF-6					

PainDetect Questionnaire Raw Score					
Nociceptive Pain					
Possible Neuropathic Pain					
Neuropathic pain					
Missing					
PHQ-2 Raw Score					
HEAL Positive Outlook Raw Score					
PROMIS-Cognitive Function - Abilities 2a Raw Score					
General Sensory Sensitivity Score - External					
General Sensory Sensitivity Score - Interoception					
General Sensory Sensitivity Score - Total					
Self-reported Sleep Duration in the Past Month (hours)					
Symptom Severity Index					
Widespread Pain Raw Score					
Stomach Pain in the Past 4 Weeks					
Not bothered at all					
Bothered a little					
Bothered a lot					
Missing					
Headaches in the Past 4 Weeks					
Not bothered at all					
Bothered a little					
Bothered a lot					
Missing					
Radiating Pain to Buttock/Thigh, Past 2 Weeks					
Yes					
No					
Not Sure					
Missing					

Radiating Pain to Below Knee, Past 2 Weeks					
Yes					
No					
Not Sure					
Missing					
Social Determinants of Health: Transportation Needs					
Yes					
No					
Missing					
Social Determinants of Health: Healthcare Needs					
Yes					
No					
Missing					
Social Determinants of Health: Food Insecurity					
Often True					
Sometimes True					
Never True					
Missing					
Social Determinants of Health: Food Money					
Often True					
Sometimes True					
Never True					
Missing					
Social Determinants of Health: Utilities					
Yes					
No					
Missing					
Social Determinants of Health: Stable Housing					
Yes					
No					
Missing					
Social Determinants of Health: Emotional Support					

Yes					
No					
Missing					
Social Determinants of Health: Number of Close Friends					
0					
1-2					
3-5					
6-10					
More than 10					
Missing					
Perceived Discrimination: Race/Ethnicity					
Never					
Rarely					
Sometimes					
Often to Almost Always					
Not sure					
Missing					
Perceived Discrimination: Orientation/Gender Identity					
Never					
Rarely					
Sometimes					
Often to Almost Always					
Not sure					
Missing					

[1]Discrete variable rows display count and percent, N (%). Continuous variable rows display mean and standard deviation, XX.X (YY.Y).
[2]Sex assigned at birth is defined as the most commonly reported identity collected from a participant across all visits.
[3]Counts retired, unemployed, and other participants who are not currently in the labor force.
[4]Counts only include respondents who have answered either 'Yes, one operation' or 'Yes, more than one operation' when asked about prior low-back operations.
[5]Participants who self-reported male sex at birth were asked about the frequency of >4 alcoholic drinks, while participants who self-reported female sex at birth were asked about the frequency of >3 alcoholic drinks.

Appendix 3 Table 1: Clinical Trials Adverse Event Reporting Table All-Cause Mortality

Population: All Participants Randomized to Stage 1 (N=805); Including All Participants Randomized to Stage 2 (N=683)

All-Cause Mortality														
	Stage 1 ACT (N=xx)	**Stage 1 Duloxetine (N=xx)**	**Stage 1 EBEM (N=xx)**	**Stage 1 ESC (N=xx)**	**Stage 2 ACT (N=xx)**	**Stage 2 Duloxetine (N=xx)**	**Stage 2 EBEM (N=xx)**	**Stage 2 ESC (N=xx)**	**Stage 2 ACT and Duloxetine (N=xx)**	**Stage 2 ACT and EBEM (N=xx)**	**Stage 2 ACT and ESC (N=xx)**	**Stage 2 Duloxetine and EBEM (N=xx)**	**Stage 2 Duloxetine and ESC (N=xx)**	**Stage 2 EBEM and ESC (N=xx)**
	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)	Affected / at Risk (%)
Total	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)	x/N (x.x%)

Appendix 3 Table 2: Clinical Trials Adverse Event Reporting Table Serious Adverse Events

Population: All Participants Randomized to Stage 1 (N=805); Including All Participants Randomized to Stage 2 (N=683)

Serious Adverse Events									
System Organ Class	**Preferred Term**	**Stage 1 ACT (N=xx)**		**Stage 1 Duloxetine (N=xx)**		**Stage 1 EBEM (N=xx)**		**Stage 1 ESC (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y

Serious Adverse Events									
System Organ Class	**Preferred Term**	**Stage 2 ACT (N=xx)**		**Stage 2 Duloxetine (N=xx)**		**Stage 2 EBEM (N=xx)**		**Stage 2 ESC (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y

Serious Adverse Events									
System Organ Class	**Preferred Term**	**Stage 2 ACT and Duloxetine (N=xx)**		**Stage 2 ACT and EBEM (N=xx)**		**Stage 2 ACT and ESC (N=xx)**		**Stage 2 Duloxetine and EBEM (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events

	Total	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y

Serious Adverse Events					
System Organ Class	**Preferred Term**	**Stage 2 Duloxetine and ESC(N=xx)**		**Stage 2 EBEM and ESC (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y

Appendix 3 Table 3: Clinical Trials Adverse Event Reporting Table Other (Not Including Serious) Adverse Events Over a Threshold of 5%

Population: All Participants Randomized to Stage 1 (N=805); Including All Participants Randomized to Stage 2 (N=683)

Other (Not Including Serious) Adverse Events									
Frequency Threshold for Reporting Other Adverse Events		5%							
System Organ Class	**Preferred Term**	**Stage 1 ACT (N=xx)**		**Stage 1 Duloxetine (N=xx)**		**Stage 1 EBEM (N=xx)**		**Stage 1 ESC (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y

Other (Not Including Serious) Adverse Events									
Frequency Threshold for Reporting Other Adverse Events		5%							
System Organ Class	**Preferred Term**	**Stage 2 ACT (N=xx)**		**Stage 2 Duloxetine (N=xx)**		**Stage 2 EBEM (N=xx)**		**Stage 2 ESC (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y

Other (Not Including Serious) Adverse Events									
Frequency Threshold for Reporting Other Adverse Events		5%							
System Organ Class	**Preferred Term**	**Stage 2 ACT and Duloxetine (N=xx)**		**Stage 2 ACT and EBEM (N=xx)**		**Stage 2 ACT and ESC (N=xx)**		**Stage 2 Duloxetine and EBEM (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y	x/N (x.x%)	y

Other (Not Including Serious) Adverse Events					
Frequency Threshold for Reporting Other Adverse Events		5%			
System Organ Class	**Preferred Term**	**Stage 2 Duloxetine and ESC(N=xx)**		**Stage 2 EBEM and ESC (N=xx)**	
		Affected / at Risk (%)	# Events	Affected / at Risk (%)	# Events
	Total	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A1**	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class A	**Preferred Term A2**	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B1**	x/N (x.x%)	y	x/N (x.x%)	y
System Organ Class B	**Preferred Term B2**	x/N (x.x%)	y	x/N (x.x%)	y