Protocol Title:

Randomized Clinical Trial of Fibromyalgia Integrative Training (FIT Teens) for adolescents with Juvenile Fibromyalgia – FIT Teens Study (Umbrella)

NCT Number:

NCT03268421

Document Date:

26 November 2024

Statistical Analysis Plan for
Multi-site Randomized Clinical Trial of FIT Teens for Juvenile Fibromyalgia

S. Kashikar-Zuck (PI), J. Peugh (Biostatistician)
R. Ittenbach (Director, DCC)
CCHMC IRB Number: 2015-6993

Grant Number: R01AR070474

Version Date: November 2024 (Revised Dec 2020, Mar 2023, July 2023, Mar 2024)

1.0 Introduction

Juvenile-onset fibromyalgia (JFM) is a chronic, debilitating pain condition that typically persists into adulthood for the majority of patients(1). Whereas medications offer limited and short-term symptom relief for JFM, prior research has demonstrated that cognitive-behavioral therapy (CBT) is safe, effective and durable in reducing functional disability and depressive symptoms in adolescents with this condition (2). However, 60% of patients receiving CBT did not show *clinically significant* improvement in functional disability, and pain levels remained in the moderate range despite being reduced overall (3). Objectively measured sedentary activity also did not significantly improve with CBT(4). Incorporation of a physical exercise component emerged as a logical next step to enhance CBT, yet regular participation in any physical activity has been shown to be difficult to initiate and maintain in FM patients.

Innovative features of recently developed interventions have included neuromuscular training designed to limit delayed muscle soreness as well as seamless integration with CBT to enhance psychological coping skills, decreased fear of movement as well as increased physical activity participation. Preliminary results have offered patient engagement, lack of adverse effects and very promising early results (5) indicating this treatment to have even stronger effects on disability and pain outcomes than CBT alone. The purpose of this randomized clinical trial is to test whether the FIT Teens intervention is more effective than CBT alone or graded aerobic exercise (GAE) alone and whether treatment effects are sustainable over 1-year follow-up.

2.0 Study Hypotheses

Primary Aims

AIM 1: Testing the effectiveness of FIT Teens in reducing functional disability (primary outcome) compared to other treatment arms (CBT alone, GAE alone)

- **H1a**: The FIT Teens group will show significantly greater reduction in functional disability at 3-month follow-up compared to CBT
- **H1b**: The FIT Teens group will show significantly greater reduction in functional disability at 3-month follow-up compared to GAE

AIM 2: Testing whether reductions in functional disability in the FIT Teens group are maintained at lower levels than CBT alone or GAE alone over time.

- **H2a**: Functional disability in the FIT Teens group will remain significantly lower than CBT at 6-, 9- and 12-month follow-up.
- **H2b**: Functional disability in the FIT Teens group will remain significantly lower than GAE at 6-, 9- and 12-month follow-up.

AIM 3: Testing whether more patients who receive FIT Teens achieve clinically meaningful improvement in functional disability compared to those who receive CBT and GAE.

- **H3a**: A significantly greater proportion of the FIT Teens group will achieve clinically meaningful reduction (i.e., a decrease in at least 7.8 points) in functional disability at 3-month follow-up than CBT.
- **H3b**: A significantly greater proportion of the FIT Teens group will achieve clinically meaningful reduction (i.e., a decrease in at least 7.8 points) in functional disability at 3-month follow-up than GAE.

Secondary Aim

Aim 4: To test whether the combined FIT Teens intervention is more effective in reducing pain intensity.

- **H4a**: The FIT Teens group will show significantly greater reduction in pain intensity at 3-month follow-up compared to CBT.
- **H4b**: The FIT Teens group will show significantly greater reduction in pain intensity at 3-month follow-up compared to GAE.
- **H4c**: Pain intensity in the FIT Teens group will remain significantly lower than CBT over time (6-, 9- and 12-month follow-up).
- **H4d**: Pain intensity in the FIT Teens group will remain significantly lower than GAE over time (6-, 9- and 12-month follow-up).

3.0 Study Design

This R01 study is a multi-site, 3-arm, prospective, randomized clinical trial that will test the efficacy of the FIT Teens intervention as compared with two established treatments, CBT and GAE, with respect to functional disability and pain among adolescents with fibromyalgia. After a baseline assessment, participants will be randomized in groups of 3-6 participants (defined as a "cohort") to receive a 16-session group-based intervention based on their treatment assignment (CBT, FIT or GAE). Longitudinal data will be collected at six time points (Baseline [BL], Post-Treatment [Post], 3-, 6-, 9-, and 12-months follow-up) at each of seven clinical sites (Cincinnati, OH; Columbus, OH; Hartford, CT; Boston, MA; Indianapolis, IN; Kansas City, MO; & Toronto, CA) over a 54-month period. The primary outcome of interest of this study is ***functional disability***, while *pain intensity* serves as a close secondary outcome.

The COVID-19 pandemic resulted in restrictions which precluded in-person delivery of group-based interventions; therefore a remote-delivery platform for all treatment arms was used as of July 2020. Trial included two alternate formats of treatment delivery (IP- in person) and (RD – remote delivery).

4.0 Power and Sample Size Estimates

(Note: the following power calculation is an amended version from the original power calculation and based on the R01 Ancillary Funded Award in February 2023 which reduced the scope of the RCT to accommodate delays due to the COVID19 pandemic)

The primary outcome for this RCT is reduction in functional disability. Power was determined for the ability to detect whether FIT Teens results in significantly greater reduction in FDI scores for Aims 1a, 1b, and 1c. Assumptions for power calculations were based on previous research showing statistically significant decreases in disability resulting from CBT (Cohen's d = .52) and GAE (Cohen's d = .40) interventions (6,7) and our pilot studies showing stronger post-treatment effect sizes for FIT Teens (d = .95 – 1.22). Power estimates were calculated with more

conservative effect sizes for FIT Teens (i.e., d = .90). All power calculations were based on the assumption of 20% attrition and proper handling of missing data.

Aim 1a. Differences between groups at the 3-month primary endpoint.
Power was calculated via the external Monte Carlo simulation capabilities in Mplus in two steps. First, 5000 dataset replications of hypothetical FDI scores were generated in a multiple group SEM format assuming: 1) standardization of FDI scores, 2) no differences in the three groups at baseline due to randomization (d = 0), 3) group differences in FDI scores at 3 months, consistent with effect sizes from prior studies as follows: GAE (d = 0.40), CBT (d = 0.52), & FIT (d estimated more conservatively at 0.90), and 4) N = *315* (n = 105 per group) available for analysis assuming proper missing data handling. Second, the 5000 Monte Carlo replications were then analyzed using a longitudinal SEM assuming linear trend (i.e., slope loadings coded 0, 2, & 3) with dummy-coded CBT & GAE groups (FIT = reference). Results showed power > 0.80 if the standardized 'slope on group' coefficient for GAE is β > 0.12 and power = 0.79 for CBT assuming proper handling of both cluster and site nesting.

Aim 1b. Maintenance of treatment gains over follow-up.
Power was again calculated via the external Monte Carlo simulation capabilities in Mplus in two steps. First, 5000 dataset replications of hypothetical FDI scores were generated in a multiple group SEM format assuming: 1) standardization of FDI scores, and 2) maintained differences between the three groups at 6-, 9-, & 12-months as follows: GAE (d = 0.40), CBT (d = 0.52), & FIT (d = 0.90), and 3) N = 315 (n = 105 per group) available for analysis assuming proper missing data handling. Second, the 5000 Monte Carlo replications were then analyzed using a longitudinal SEM assuming an intercept-only model (i.e., slope fixed and random effects both = 0) with dummy-coded CBT & GAE groups (FIT = reference). Results showed power **> 0.80** if the standardized 'intercept on group' coefficient for either CBT or GAE is β > 0.38 assuming proper handling of both cluster and site nesting.

Aim 1c. Proportion of patients achieving clinically meaningful change.
We anticipate that the FIT Teens intervention will result in a greater proportion of patients achieving the binary outcome of clinically meaningful reduction in FDI compared to CBT and GAE. This difference is expected to be between 15%-20% based on our pilot study of FIT teens and our previous trial of CBT showing ~55% of the FIT group and 35-40% of CBT participants achieving clinically meaningful change. The proportion of GAE patients achieving this binary outcome is expected to be similar to CBT. Power was calculated via G*Power3 assuming proper handling of missing data and n = 105 within the three arms available for analysis. Results showed power will be 0.79 for either the FIT vs. CBT or FIT vs. GAE comparison if difference between two independent proportions is at least 17% or greater (False Discovery Rate Type-1 error control will be used to evaluate results from both tests). A difference in proportion of ≥17% between FIT Teens and CBT or GAE will provide useful information for patients and providers about the relative efficacy of the interventions in achieving clinically meaningful improvement in functioning.

5.0 Randomization

Participants will be randomized into FIT, CBT or GAE groups based upon a randomization schedule maintained by the study biostatistician. A randomization process *by group* will be used. In other words, after a sufficient number of approximately 4 to 6 patients have been screened at a site and found to be eligible, the biostatistician will inform the therapist of the next group assignment. Randomization will be stratified by site in order to ensure approximately equal proportions of patients from each site in each of the 3 arms. Group assignments will be masked to the PI and study assessors. The study biostatistician (Peugh) will monitor the equivalence of sample size in each treatment arm and at study mid-point will make any adjustments to the randomization schedule in consultation with the DSMB to ensure approximately equal group sizes by the end of the study.

6.0 Derivations and Definitions

Variable Definitions

Demographics and Study Status

Variable	Descriptor	Derivation (if applicable)
Site	Institution Name	
SiteID (derived)	Site ID Number	IF Site=Cincinnati Children's Hospital Medical Center, 10; IF Site= Children's Mercy Hospitals and Clinics, 30; IF Site=Connecticut Children's Medical Center, 40; IF Site= Nationwide Children's Hospital, 50; IF Site= The Hospital for Sick Children, 60; IF Site= Boston Children's Hospital, 70; IF Site= Indianapolis Riley Children's Hospital, 80
Subject	Study ID [2-digit site ID]-[3-digit subject ID]	
mGROUP	Masked Group Assignment (A, B, C)	(From Randomization Record) (mGROUP to be edited to GROUP after primarily analysis complete)
GROUP	Group Assignment (CBT, FIT, GAE)	(From Randomization Record)
FORMAT	Treatment Delivery Format (In-person=1, Remote=2)	(From Randomization Record)
COHORTID_STD	Cohort ID (01-18)	
RFENROLLDTE_RAW	Date of Enrollment (Signed Consent)	
BRTHDAT_RAW	Participant's date of birth	
SEX_STD	Participant's sex 1=male; 2=female; 99 = other	
RACE_STD	Participant's race	**Variable Labels:** 1=More than one race; 2=Caucasian/White; 3=African American/Black; 4=Asian; 5=Native Hawaiian/Other Pacific Islander; 6= American Indian/Alaskan Native
ETHNIC_STD	Participant's ethnicity	**Variable Labels:** 1=Hispanic; 2=Non-Hispanic
INCOME_STD	Family Income	**Variable Labels:** 1=Less than $24,999; 2=$25,000 - $49,999;

Variable	Descriptor	Derivation (if applicable)
		3=$50,000 - $74,999; 4=$75,000 - $99,999; 5=$100,000 - $124,999; 6=$125,000 - $149,999; 7=More than $150,000; 8=Prefer not to answer; 99=Unknown
INSRNCE_STD	Insurance Status	**Variable Labels:** 1=Private Health Insurance; 2=Medicare; 3=Medicaid; 4=Military Health Care; 5=State Specific; 6= Indian Health Services; 7=Non-Us; 8=Other; 9 = None
CRGVR1_STD	Caregiver 1	**Variable Labels:** 1=Mother; 2=Father; 3=Legal Guardian
EDU1_STD	Caregiver 1 level of education	**Variable Labels:** 1=Less than High School; 2=High School/GED; 3=Partial College or Trade School; 4=College Graduate; 5=Post Graduate Degree
CRGVR2_STD	Caregiver 2	**Variable Labels:** 1=Mother; 2=Father; 3=Legal Guardian
EDU2_STD	Caregiver 2 level of education	**Variable Labels:** 1=Less than High School; 2=High School/GED; 3=Partial College or Trade School; 4=College Graduate; 5=Post Graduate Degree
EDUENRGRD_STD	Participant's grade enrolled	**Variable Labels:** 1=5th, 2=6th, 3=7th, 4=8th, 5=9th, 6=10th, 7=11th, 8=12th

Assessment Time Point

Variable	Descriptor	Derivation
TIMEPOINT	Assessment Visit number	IF FolderName=T1 - (Baseline); IF FolderName=T2 - (Tx Session 8 - Mid-Tx Check-In); IF FolderName=T3 - (Post-Tx Assessment); IF FolderName=T4 - (3-month Follow-Up); IF FolderName=T5 - (6-month Follow-Up); IF FolderName=T6 - (9-month Follow-Up);

		IF FolderName=T7 - (12-month Follow-Up)

Derived Variable Definitions

Demographics

Variable	Descriptor	Derivation (if applicable)
T1AGE (derived)	Participant age at enrollment in years NOTE: RFENROLLDTE is pulled from the Randomization Form (T1 folder); BIRTHDAT is pulled from the Demographics form (pre-screen folder)	(RFENROLLDTE_RAW – BIRTHDAT_RAW)/365.25

Primary Outcome: Functional Disability T1-T7

Variable	Descriptor	Derivation (if applicable)
FDITOT (derived)	Functional Disability Inventory score	Σ(FDIBATH_STD, FDISTAIRS_STD, FDIFRIEND_STD, FDICHORES_STD, FDIEATING_STD, FDIREST_STD, FDIRIDING_STD, FDISCHOOL_STD, FDIACTIVITIES_STD, FDIHOMEWORK_STD, FDIWATCHING_STD, FDIWALK_STD, FDIRUN_STD, FDISHOP_STD, FDISLEEP_STD)

Secondary Outcome: Pain Intensity Rating **T1-T7**

Variable	Descriptor	Derivation (if applicable)
PRFAVGPN_RAW	Average Pain Diary VAS rating	
FDIMPSCR_RAW	Retrospective Pain VAS Rating	

Other Variables T1, T3-T7

Variable	Descriptor	Derivation (if applicable)
CDITOT (derived)	Children's Depression Inventory 2 Raw score	COMPUTE CDI2ITEM1_STD = CDI2ITEM1_STD -1. COMPUTE CDI2ITEM2_STD = CDI2ITEM2_STD -1. COMPUTE CDI2ITEM3_STD = CDI2ITEM3_STD -1. COMPUTE CDI2ITEM4_STD= CDI2ITEM4_STD -1. COMPUTE CDI2ITEM5_STD= CDI2ITEM5_STD -1. COMPUTE CDI2ITEM6_STD= CDI2ITEM6_STD -1. COMPUTE CDI2ITEM7_STD= CDI2ITEM7_STD -1.

Variable	Descriptor	Derivation (if applicable)
		COMPUTE CDI2ITEM8_STD= CDI2ITEM8_STD -1. COMPUTE CDI2ITEM9_STD= CDI2ITEM9_STD -1. COMPUTE CDI2ITEM10_STD= CDI2ITEM10_STD -1. COMPUTE CDI2ITEM11_STD = CDI2ITEM11_STD -1. COMPUTE CDI2ITEM12_STD = CDI2ITEM12_STD -1. COMPUTE CDI2ITEM13_STD = CDI2ITEM13_STD -1. COMPUTE CDI2ITEM14_STD = CDI2ITEM14_STD -1. COMPUTE CDI2ITEM15_STD = CDI2ITEM15_STD -1. COMPUTE CDI2ITEM16_STD = CDI2ITEM16_STD -1. COMPUTE CDI2ITEM17_STD = CDI2ITEM17_STD -1. COMPUTE CDI2ITEM18_STD = CDI2ITEM18_STD -1. COMPUTE CDI2ITEM19_STD = CDI2ITEM19_STD -1. COMPUTE CDI2ITEM20_STD = CDI2ITEM20_STD -1. COMPUTE CDI2ITEM21_STD = CDI2ITEM21_STD -1. COMPUTE CDI2ITEM22_STD = CDI2ITEM22_STD -1. COMPUTE CDI2ITEM23_STD = CDI2ITEM23_STD -1. COMPUTE CDI2ITEM24_STD = CDI2ITEM24_STD -1. COMPUTE CDI2ITEM25_STD = CDI2ITEM25_STD -1. COMPUTE CDI2ITEM26_STD = CDI2ITEM26_STD -1. COMPUTE CDI2ITEM27_STD = CDI2ITEM27_STD -1. COMPUTE CDI2ITEM28_STD = CDI2ITEM28_STD -1. EXECUTE. RECODE CDI2ITEM2_STD CDI2ITEM6_STD CDI2ITEM7_STD CDI2ITEM9_STD CDI2ITEM10_STD CDI2ITEM12_STD CDI2ITEM14_STD CDI2ITEM15_STD CDI2ITEM17_STD CDI2ITEM20_STD CDI2ITEM23_STD CDI2ITEM24_STD CDI2ITEM26_STD CDI2ITEM27_STD (2=0) (1=1) (0=2). EXECUTE. COMPUTE CDI_sum = CDI2ITEM1_STD+CDI2ITEM2_STD+CDI2ITEM3_STD+CDI2ITEM4_STD+CDI2ITEM5_STD+CDI2ITEM6_STD+CDI2ITEM7_STD+CDI2ITEM8_STD+CDI2ITEM9_STD+CDI2ITEM10_STD+CDI2ITEM11_STD+CDI2ITEM12_STD+CDI2ITEM13_STD+CDI2ITEM1

Variable	Descriptor	Derivation (if applicable)
		4_STD+CDI2ITEM15_STD+CDI2ITEM16_STD+CDI2ITEM17_STD+CDI2ITEM18_STD+CDI2ITEM19_STD+CDI2ITEM20_STD+CDI2ITEM21_STD+CDI2ITEM22_STD+CDI2ITEM23_STD+CDI2ITEM24_STD+CDI2ITEM25_STD+CDI2ITEM26_STD+CDI2ITEM27_STD+CDI2ITEM28_STD. EXECUTE.
PCSTOT (derived)	Pain Catastrophizing Scale score	∑(PCS1_STD,…,PCS13_STD)
PCQTOT (derived)	Pain Coping Efficacy score	RECODE PCQDEAL_STD (1=5) (2=4) (3=3) (4=2) (5=1) EXECUTE. COMPUTE PCQTOT = PCQCHANGE_STD + PCQDEAL + PCQMOOD_STD. EXECUTE.
TSKTOT (derived)	Tampa Scale of Kinesiophobia score	∑(TSKHURT_STD,TSKPUSH_STD,TSKWRONG_STD,TSKCONDITION_STD,TSKRISK_STD,TSKINJURED_STD,TSKCAREFUL_STD, TSKDANGEROUS_STD,TSKEXERCISE_STD, TSKNORMAL,TSKPAIN_STD)
PAININT (derived)	PROMIS Pain Interference	∑(PPISLEEP_STD, PPIANGRY_STD, PPISCHOOLWK_STD, PPIATTENTION_STD, PPIRUN_STD, PPIWALK_STD, PPIFUN_STD, PPISTAND_STD)
PAININTt (derived)	PROMIS Pain Interference t-score	IF PAININT = 8, 34.0 IF PAININT = 9, 38.7 IF PAININT = 10, 40.6 IF PAININT = 11, 42.7 IF PAININT = 12, 44.3 IF PAININT = 13, 45.8 IF PAININT = 14, 47.1 IF PAININT = 15, 48.4 IF PAININT = 16, 49.5 IF PAININT = 17, 50.6 IF PAININT = 18, 51.7 IF PAININT = 19, 52.7 IF PAININT = 20, 53.7 IF PAININT = 21, 54.7 IF PAININT = 22, 55.7

Variable	Descriptor	Derivation (if applicable)
		IF PAININT = 23, 56.6 IF PAININT = 24, 57.6 IF PAININT = 25, 58.5 IF PAININT = 26, 59.5 IF PAININT = 27, 60.4 IF PAININT = 28, 61.4 IF PAININT = 29, 62.4 IF PAININT = 30, 63.4 IF PAININT = 31, 64.4 IF PAININT = 32, 65.4 IF PAININT = 33, 66.5 IF PAININT = 34, 67.6 IF PAININT = 35, 68.8 IF PAININT = 36, 70.1 IF PAININT = 37, 71.5 IF PAININT = 38, 73.2 IF PAININT = 39, 75.0 IF PAININT = 40, 78.0
FATIG (derived)	PROMIS Fatigue	∑(PFFRIENDS_STD, PFWEAK_STD, PFEASILY_STD, PFSCHOOLWK_STD, PFFINISHING_STD, PFSTARTING_STD, PFATTENTION_STD, PFSPORTS_STD, PFTHINGS_STD, PFENJOY_STD)
FATIGt (derived)	PROMIS Fatigue t-score	IF FATIG = 10, 30.3 IF FATIG = 11, 34.3 IF FATIG = 12, 36.9 IF FATIG = 13, 39.0 IF FATIG = 14, 40.9 IF FATIG = 15, 42.5 IF FATIG = 16, 44.0 IF FATIG = 17, 45.4 IF FATIG = 18, 46.7 IF FATIG = 19, 47.9 IF FATIG = 20, 49.1 IF FATIG = 21, 50.2 IF FATIG = 22, 51.3 IF FATIG = 23, 52.4 IF FATIG = 24, 53.5 IF FATIG = 25, 54.5 IF FATIG = 26, 55.6 IF FATIG = 27, 56.6 IF FATIG = 28, 57.6 IF FATIG = 29, 58.6 IF FATIG = 30, 59.6 IF FATIG = 31, 60.6

Variable	Descriptor	Derivation (if applicable)
		IF FATIG = 32, 61.6 IF FATIG = 33, 62.6 IF FATIG = 34, 63.6 IF FATIG = 35, 64.6 IF FATIG = 36, 65.6 IF FATIG = 37, 66.7 IF FATIG = 38, 67.7 IF FATIG = 39, 68.7 IF FATIG = 40, 69.8 IF FATIG = 41, 70.9 IF FATIG = 42, 72.0 IF FATIG = 43, 73.2 IF FATIG = 44, 74.4 IF FATIG = 45, 75.7 IF FATIG = 46, 77.0 IF FATIG = 47, 78.5 IF FATIG = 48, 80.2 IF FATIG = 49, 82.0 IF FATIG = 50, 84.0
PAINBEH (derived)	PROMIS Pain Behavior	∑(PPBFACE_STD, PPBMEDICINE_STD, PPBTALKED_STD, PPBMOVED_STD, PPBPROTECTED_STD, PPBSTOP_STD, PPBHELP_STD, PPBLAYDOWN_STD)
PAINBEHt (derived)	PROMIS Pain Behavior t-score	IF PAINBEH = 8, 20.0 IF PAINBEH = 9, 28.4 IF PAINBEH = 10, 30.7 IF PAINBEH = 11, 32.6 IF PAINBEH = 12, 34.2 IF PAINBEH = 13, 35.7 IF PAINBEH = 14, 37.1 IF PAINBEH = 15, 38.5 IF PAINBEH = 16, 39.8 IF PAINBEH = 17, 40.9 IF PAINBEH = 18, 41.9 IF PAINBEH = 19, 42.9 IF PAINBEH = 20, 43.9 IF PAINBEH = 21, 44.8 IF PAINBEH = 22, 45.7 IF PAINBEH = 23, 46.5 IF PAINBEH = 24, 47.3 IF PAINBEH = 25, 48.1 IF PAINBEH = 26, 48.8 IF PAINBEH = 27, 49.6 IF PAINBEH = 28, 50.3 IF PAINBEH = 29, 51.0

Variable	Descriptor	Derivation (if applicable)
		IF PAINBEH = 30, 51.8 IF PAINBEH = 31, 52.5 IF PAINBEH = 32, 53.2 IF PAINBEH = 33, 53.9 IF PAINBEH = 34, 54.7 IF PAINBEH = 35, 55.4 IF PAINBEH = 36, 56.2 IF PAINBEH = 37, 57.0 IF PAINBEH = 38, 57.7 IF PAINBEH = 39, 58.5 IF PAINBEH = 40, 59.4 IF PAINBEH = 41,60.3 IF PAINBEH = 42, 61.2 IF PAINBEH = 43, 62.2 IF PAINBEH = 44, 63.3 IF PAINBEH = 45, 64.6 IF PAINBEH = 46, 66.2 IF PAINBEH = 47, 67.9 IF PAINBEH = 48, 80.0
WPITOT (derived)	Widespread Pain Index	∑(PSAQSHLDRT, PSAQSHLDLT, PSAQUARMRT, PSAQUARMLT, PSAQLARMRT, PSAQLARMLT, PSAQHIPRT, PSAQHIPLT, PSAQULEGRT, PSAQULEGLT, PSAQLLEGRT, PSAQLLEGLT, PSAQJAWRT, PSAQJAWLT, PSAQCHEST, PSAQUBACK, PSAQLBACK, PSAQNECK)
CARDSX (derived)	Cardinal Symptoms of JFM	∑(PSAQFATIGUE_STD, PSAQTIRED_STD, PSAQMEMORY_STD)
SSCHK (derived)	Somatic Symptoms Total	∑(PSAQWEAKNESS,PSAQNERVOUS,PSAQEYES, PSAQNUMB, PSAQDEPRESS, PSAQITCHING, PSQAMIGRAINE, PSAQURINATION, PSAQDIZZY, PSAQCRAMPS, PSAQAPPETITE, PSAQBREATH, PSAQCONSTIPATE, PSAQVISION, PSAQTHINK,PSAQHRTBURN, PSAQEARS, PSAQMOUTH, PSAQNAUSEA, PSAQBRUISE, PSAQTENDER, PSAQIBS, PSAQINSOMNIA, PSAQSENSITIVE)

Variable	Descriptor	Derivation (if applicable)
SOMSX (derived)	Somatic Symptoms Score 0=no symptoms 1=few symptoms 2=moderate symptoms 3=great deal of symptoms	IF SSCHK = 0, SOMSX =“0”, IF SSCHK GT 0 AND <= 5, SOMSX = “1”, IF SSCHK >= 6 & <= 9, SOMSX =“2”, IF SSCHK >=10, SOMSX =“3”
SSITOT (derived)	Symptom Score	CARDSX + SOMSX
PSAT_TOT (derived)	PSAT Total Score	∑(WPITOT, SSITOT)
FMGROUP (derived)	Juvenile Fibromyalgia Groups 0=Non-JFM 1=JFM	IF WPITOT >=7 AND SSITOT>=5, FMGROUP = “1”; or IF WPITOT >=3 AND <= 6 AND SSITOT>=9, FMGROUP = “1”; ELSE FMGROUP = “0”
CASPETOT (derived)	COVID Impact Score	∑(CASPE1_STD…CASPE16_STD)
BEIGHTON (derived)	Joint Hypermobility Score (Range: 0-9)	PERTFINGER_STD + PELTFINGER_STD + PERTELBOW_STD + PELTELBOW_STD + PERTTHUMB_STD + PELTTHUMB_STD + PERTKNEE_STD + PELTKNEE_STD + PEPALMS_STD
RSOVERALL_STD	Clinician Global Assessment Scale Range 1-10; 1=Very Poorly, 10=Very Well	
MEDGROUP	Classification of specific medication into medication class	1=antidepressants; 2=other psychotropics; 3=NSAIDS; 4=anticonvulsants; 5=musclerelaxants; 6=non-opioid analgesics
MEDGROUPYN	Participant taking any medications within medication class	1=Yes 0=No
MEDGROUPTOT	Total medications taken per medication class	
MEDTOT	Total number of medications overall	

5.0 Analysis Population

A total of 315 adolescents 12.0 to 18.0 years of age (*n* = ~105 / group for 3 IV groups) will be needed for this study. The samples used for analysis are as follows:

Sample 1. All JFM subjects with a diagnosis of fibromyalgia consistent with current American College of Rheumatology criteria, irrespective of intervention group (*n* = 315).
Sample 2. All JFM subjects receiving the FIT Teens intervention (*n* = 105).
Sample 3. All JFM subjects receiving the CBT treatment (*n* = 105).
Sample 4. All JFM subjects receiving the GAE treatment (*n* = 105).

6.0 Statistical Analyses

Analyses will be carried out on the full intent-to-treat sample (Sample 1) as the primary analysis. Data analysis will begin with a review of all relevant variables in the dataset. For continuous variables, parametric as well as nonparametric measures of central tendency, variability, and association, will be computed. Distributional properties of potential outcomes will be evaluated and tested for normality where appropriate. Those differing markedly from normality will be considered candidates for transformation or alternative modeling techniques. Our analyses will be divided into three distinct sections. First (Primary Aims), the effectiveness of the FIT Teens intervention (Sample 2) for the reduction of functional disability will be compared with the other two treatments in several ways. FIT Teens will be compared to a CBT-only group (Sample 3) and GAE group (Sample 4) at 3-month follow-up (H1a, H1b). Additionally, differences between groups will be compared over time (i.e., 6, 9, 12-month follow-up; H2a, H2b), and analyses will be conducted to determine whether more patients in the FIT Teens intervention group obtain clinically meaningful improvement at 3-months when compared to the other two treatments (H3a, H3b).

Second, (Secondary Aims), the effectiveness of the FIT Teens intervention (Sample 2) for the reduction of pain intensity will be compared with the other two treatments (Samples 3 and 4) in several ways, similar to FDI analyses, including comparisons at 3-month follow-up (H4a, H4b), as well as whether these treatments are sustained over time (i.e., 6, 9, 12-month follow-up; H4c, H4d). Last, exploratory analyses also will be conducted to examine a mechanistic model of physical activity. Unless otherwise noted, $\alpha = 0.05$ (two-sided) will serve as the criterion for statistical significance for all analyses. All data will be analyzed using the current version of MPlus.

6.1 Descriptive Statistics

Data analysis will begin with a review of all relevant variables in the dataset. For continuous variables, parametric measures of central tendency, variability, and association, will be computed. Distributional properties of potential outcomes will be evaluated and robust estimation techniques (MLR estimation, bootstrap resampling) will be used to guard against Type-1 errors that typically result from non-normal response variable data.

Descriptive statistics. Descriptive statistics (mean, median, std, min, max) will be computed for the following variables at each determined assessment time point: (with baseline only baseline only for demographics): e.g., AGE, INCOME, MEDTOT, CGA, PRFAVGPN_RAW, FDITOT, CDITOT, PCSTOT, PCQTOT, TSKTOT, FATIG,

PAINBEH, PAININT, WPITOT, CARDSX, SSITOT, SOMSX, SSCHK, CASPETOT, BEIGHTON

Frequency Counts/Percentages. Frequency counts and percentages will be computed for the following variables: GROUP, FORMAT, SEX, RACE, ETHNIC, INCOME, INSURANCE, CRGVR1, EDU1, CRGVR2, EDU2, ENGRGRD, FMGROUP

6.2 Baseline Comparisons

Baseline group comparisons will be used to test for equivalence among the three intervention groups at baseline (Samples 1, 2, 3 by GROUP)

- One-way ANOVA/Kruskall Wallis: AGE, INCOME, RSOVERALL, FDITOT, APVAS, FDITOT, CDITOT, PCSTOT, PCQTOT, TSKTOT, FATIG, PAINBEH, INTERF, WPITOT, CARDSX, SSITOT, SOMSX, SSCHK, CASPETOT, BEIGHTON
- Chi-square/Exact Test: FORMAT, SEX, RACE, ETHNIC, INCOME, INSURANCE, CRGVR1, EDU1, CRGVR2, EDU2, ENGRGRD, FMGROUP

Specific tests used will depend on distributional properties and expected cell counts. Test results with *p*-values <.10 will be considered as candidates for inclusion in the multiple covariate models in the analyses below.

6.3 Primary Analyses

This research design constitutes a four-level model: repeated longitudinal assessments (level 1) are nested within participants (level 2) who are tested in cohorts of 4-6 participants (level 3) within each of the six study sites (level 4). Prior to testing the effects of treatment group on primary and secondary outcomes, Cohort (Level 3) and Site (Level 4) variation must be addressed appropriately. Study site serves only as a means to an end regarding recruitment of the requisite sample size. Cohort (level 3) is a unique design and implementation feature of group-based sessions that likely serves several beneficial purposes: social support from similar age peers, greater engagement and participation, decreased likelihood of dropout, amongst others. However, the primary aims concern only the efficacy of FIT-TEENS versus GAE and CBT at the level of the individual participant (level 2). As such, study site (level 4) and cohort (level 3) constitute statistical “nuisances” that must be dealt with correctly with to avoid biased treatment effects and subsequent inferential errors. Specifically, primary and secondary aims analyses will use generalized estimating equations for two reasons: 1.) continuous repose variables need not be normally distributed, and 2.) the random effects structure (i.e., covariance matrix) is not of interest. Specifically, every unique combination of participant (level 2), cohort (level 3), and site (level 4) will be used to indicate non-independent (i.e., correlated) responses in the GEE analyses.

Also, as noted in Study Design (Section 3.0), with the onset of COVID 19 restrictions in March 2020, all treatment arms will be delivered in remote format. A binary indicator variable (0= In-Person delivery; 1 = Remote Delivery) will be included as a moderator in all analyses to model the effect of delivery format. Further, the need for an indicator variable by independent variable group interaction term (e.g., format X GAE, format X CBT) to adjust for any potential deviations from the randomization plan due to the change in delivery format will be tested.

H1a The FIT Teens group will show significantly greater reduction in functional disability at 3-month follow-up compared to CBT.

H1b The FIT Teens group will show significantly greater reduction in functional disability at 3-month follow-up compared to GAE.

Results will be analyzed via a longitudinal GEE approach. Groups will be dummy-coded with FIT as the reference class. It is hypothesized that significant and positive 'slope on group' coefficients for CBT and GAE indicating a significantly lower FDI for FIT vs. CBT & GAE. Non-independence of FDI scores will be addressed by specifying every unique combination of participant (level 2), cohort (level 3), and site (level 4) as indicators of non-independent (correlated) responses in the GEE syntax (i.e., in the 'Repeated =' syntax specification.

H2a Functional disability in the FIT Teens group will remain significantly lower than CBT at 6-, 9- and 12-month follow-up.

H2b Functional disability in the FIT Teens group will remain significantly lower than GAE at 6-, 9- and 12-month follow-up.

To test whether group differences on FDI scores are maintained over time, a longitudinal SEM approach will again be used. Groups will be dummy-coded with FIT as the reference class. Significant and positive 'intercept on CBT' & 'intercept on GAE' coefficients will indicate significantly lower FDI scores for FIT vs. CBT & FIT vs GAE at 6-, 9-, & 12-month assessments. Significant and positive 'slope on CBT' & 'slope on GAE' coefficients would indicate worsening FDI scores for CBT & GAE over time relative to FIT scores that have stayed the same or further improved. Nesting of clustered participants within sites will be handled with the SEM model in Mplus described above.

H3a A significantly greater proportion of the FIT Teens group will achieve clinically meaningful reduction (i.e., a decrease in at least 7.8 points) in functional disability at 3-month follow-up than CBT.

H3b A significantly greater proportion of the FIT Teens group will achieve clinically meaningful reduction (i.e., a decrease in at least 7.8 points) in functional disability at 3-month follow-up than GAE.

To test changes in the dichotomous (improved vs not improved) endpoints of functional disability from baseline to 3-month follow-up, baseline FDI scores will be subtracted from the 3-month follow-up FDI to identify those who did and did not achieve a clinically significant FDI change score. Results will be analyzed via separate difference between two independent proportions analyses for FIT vs. CBT & FIT vs. GAE testing H_0: 50% of participants in the three study arms will achieve a reliable FDI decrease versus H_A: 55% of participants in the FIT arm, and 35%-40% of participants in each of the CBT and GAE arms, will achieve a clinically significant FDI decrease.

6.4 Secondary Analyses

H4a The FIT Teens group will show significantly greater reduction in pain at 3-month follow-up compared to CBT.

H4b The FIT Teens group will show significantly greater reduction in pain at 3-month follow-up compared to GAE.

Results will be analyzed via a longitudinal SEM approach using MPlus (Version 7.4) software. Groups will be dummy-coded with pain as the reference class. It is hypothesized that significant and positive 'slope on group' coefficients for CBT and GAE indicating a significantly lower pain for FIT vs. CBT & GAE. Non-independence of pain scores within participant clusters and within sites will be addressed in three steps: 1.) declaring site as the complex clustering variable (e.g., 'Type=Complex' in Mplus), 2.) estimating a saturated patient cluster-level model (i.e., estimating all possible covariances among pain repeated measures variances at Level 2) so that, 3.) unbiased parameter estimates and significance tests can be obtained from the longitudinal SEM growth model specified and tested similar to above at the participant level.

H4c Pain in the FIT Teens group will remain significantly lower than CBT at 6-, 9- and 12-month follow-up.

H4d Pain in the FIT Teens group will remain significantly lower than GAE at 6-, 9- and 12-month follow-up.

To test whether group differences on pain scores are maintained over time, a longitudinal SEM approach will again be used. Groups will be dummy-coded with FIT as the reference class. Significant and positive 'intercept on CBT' & 'intercept on GAE' coefficients will indicate significantly lower pain scores for FIT vs. CBT & FIT vs GAE at 6-, 9-, & 12-month assessments. Significant and positive 'slope on CBT' & 'slope on GAE' coefficients would indicate worsening pain scores for CBT & GAE over time relative to FIT scores that have stayed the same or further improved. Nesting of clustered participants within sites will be handled with the SEM model in Mplus described above.

6.5 Exploratory Analyses

Mechanistic models of how treatment-related improvements in psychological (coping, fear of movement, adherence) and physical factors (objectively measured physical activity, biomechanical changes and fitness) explain changes in disability and pain outcomes across the three groups will be examined on an **exploratory basis**. Multiple path mediation analyses. 1) Selected indirect pathways between psychological coping skills, decreased fear of activity, and improved biomechanics (predictors), physical activity engagement and physical fitness (mediators), and functional disability and pain intensity (outcomes) will be tested for statistical significance within each of the three (FIT, CBT & GAE) groups, and 2) The magnitudes of the indirect effects will be tested for statistically significant differences between the three treatment groups. The False Discovery Rate will be used to maintain the nominal Type-1 error rate for all indirect effect tests.

6.5.1 Missing Data Planning and Analyses

Procedures are in place to minimize missing and inappropriate data (i.e., automated query resolution procedures in Medi-Data Rave®). Missing data will be handled via model-based 3-level multiple imputation in Mplus (version 8.10) as follows. Longitudinal growth curve models will be specified at level 1 (i.e. repeated measures [level 1] nested within participants [level 2]), level 2 (participants within cohorts), and at level 3 (cohorts within assessment sites). However, the effects of IV group random assignment will only be estimated within the level 1 growth curve model. The default potential scale reduction (PSR) value of 1.05 that determines convergence will be over-ridden and PSR = 1.01 will define convergence owing to model complexity.

9.0 References

1. Kashikar-Zuck S, Cunningham N, Sil S, Bromberg MH, Lynch-Jordan AM, Strotman D, et al. Long-term outcomes of adolescents with juvenile-onset fibromyalgia in early adulthood. Pediatrics. 2014;133(3):e592-e600.
2. Kashikar-Zuck S, Ting TV, Arnold LM, Bean J, Powers SW, Graham TB, et al. Cognitive behavioral therapy for the treatment of juvenile fibromyalgia: A multisite, single-blind, randomized, controlled clinical trial. Arthritis Rheum. 2012;64(1):297-305.
3. Sil S, Arnold LM, Lynch-Jordan A, Ting TV, Peugh J, Cunningham N, et al. Identifying treatment responders and predictors of improvement after cognitive-behavioral therapy for juvenile fibromyalgia. PAIN®. 2014.
4. Kashikar-Zuck S, Flowers SR, Strotman D, Sil S, Ting TV, Schikler KN. Physical activity monitoring in adolescents with juvenile fibromyalgia: findings from a clinical trial of cognitive-behavioral therapy. Arthritis Care Res (Hoboken). 2013;65(3):398-405.
5. Kashikar-Zuck S, Tran ST, Barnett K, Bromberg MH, Strotman D, Sil S, et al. A Qualitative Examination of a New Combined Cognitive-behavioral and Neuromuscular Training Intervention for Juvenile Fibromyalgia. Clin J Pain. 2015;32(1):70-81.
6. Kashikar-Zuck S, Ting TV, Arnold LM, et al. Cognitive behavioral therapy for the treatment of juvenile fibromyalgia: A multisite, single-blind, randomized, controlled clinical trial. *Arthritis Rheum.* Jan 2012;64(1):297-305.
7. American Pain Society. *Guideline for the management of fibromyalgia syndrome pain in adults and children.* Glenview, IL: American Pain Society; 2005.
8. Sil S, Arnold LM, Lynch-Jordan A, et al. Identifying treatment responders and predictors of improvement after cognitive-behavioral therapy for juvenile fibromyalgia. *Pain.* 2014; 155: 1206-1212

10.0 Tables – See below

Table 1. Baseline subject characteristics by treatment group and overall

Variable	FIT Teens (n=X) M (SD), %	CBT (n=Y) M (SD), %	GAE (n=Z) M (SD), %	Overall (N=XYZ) M (SD), %
Child Characteristics				
Child Age				
Female (%)				
Ethnicity (%)				
Hispanic				
Non-Hispanic				
Race (%)				
More than one				
Caucasian				
African American				
Asian				
Native Hawaiian/ Pacific Islander				
American Indian/ Native Alaskan				
Grade Level (%)				
5th				
6th				
7th				
8th				
9th				
10th				
11th				
12th				

Table 2. Baseline parent and family characteristics by treatment group and overall

Variable	FIT Teens (*n*=X)	CBT (*n*=Y)	GAE (*n*=Z)	Overall (N=XYZ)
Income				
<$24,999				
$25,000-49,999				
$50,000-74,999				
$75,000-99,999				
$100,000-124,999				
$125,000-149,000				
> $150,000				
Not Reported				
Caregiver				
Mother				
Father				
Legal guardian				
Caregiver Education				
< High school				
High school/GED				
Partial College/ Trade School				
College Degree				
Post Graduate				
Insurance				
Private				
Medicare				
Medicaid				
Military				
State Specific				
Indian Health				
Non-Us				
Other				
None				

Table 3. Outcome scores at Baseline, Post-Treatment, and Primary (3-month) Endpoint

	Baseline					Post-Treatment					3-Month Endpoint				
			95% CI					*95% CI*					*95% CI*		
	n	***Adj. M***	***LL***	***UL***	***SE_M***	***n***	***Adj. M***	***LL***	***UL***	***SE_M***	***n***	***Adj. M***	***LL***	***UL***	***SE_M***
Functional Disability															
Fit Teens															
CBT															
GAE															
Pain															
Fit Teens															
CBT															
GAE															
Child Depression															
Fit Teens															
CBT															
GAE															
Pain Catastrophizing															
Fit Teens															
CBT															
GAE															
Fear of Movement															
Fit Teens															
CBT															
GAE															
Pain Coping															
Fit Teens															
CBT															
GAE															

95% CI = 95% Confidence Interval; Adj. M = Mean values adjusted for covariates; LL = Lower Limit for 95% Confidence Interval; UL = Upper Limit for 95% Confidence Interval; SE_M = Standard Error of the Adjusted Mean

Table 4. Outcome scores at Follow-Up Time points

	6-Month					9-Month					12-Month				
			95% CI					*95% CI*					*95% CI*		
	n	***Adj. M***	***LL***	***UL***	***SE_M***	***n***	***Adj. M***	***LL***	***UL***	***SE_M***	***n***	***Adj. M***	***LL***	***UL***	***SE_M***
Functional Disability															
Fit Teens															
CBT															
GAE															
Pain															
Fit Teens															
CBT															
GAE															
Child Depression															
Fit Teens															
CBT															
GAE															
Pain Catastrophizing															
Fit Teens															
CBT															
GAE															
Fear of Movement															
Fit Teens															
CBT															
GAE															
Pain Coping															
Fit Teens															
CBT															
GAE															

95% CI = 95% Confidence Interval; Adj. M = Mean values adjusted for covariates; LL = Lower Limit for 95% Confidence Interval; UL = Upper Limit for 95% Confidence Interval; SE_M = Standard Error of the Adjusted Mean